Analysis of different sets of spin-adapted substitution operators in open-shell coupled cluster theory

Nils Herrmann and Michael Hanrath

Institute for Theoretical Chemistry, University of Cologne, Cologne, Germany

ABSTRACT

In spin-adapted open-shell coupled cluster (CC) theory, the choice of spin-free spatial substitution operators is generally not unique. Due to an increasing linear dependence of the cluster operator (with increasing substitution level), the options to span identical linear spaces increase rapidly. In this work, several sets of non-orthogonal as well as orthogonal spin-adapted substitution operators are generated and used in consecutive configuration interaction (CI) and CC calculations. All (full) operator sets were generated to span the same linear space. The results are analysed in terms of the produced wave function quality and the amount of recovered correlation energy w.r.t. full CI (FCI). In particular, the influence of different amounts of spectators, orthogonality as well as spin incompleteness was investigated. It was found that CC calculations involving fewer spectators lead to more accurate results in general. Here, correlation energy differences of up to 0.32% for minimal to maximal spectating sets were obtained. As expected, all conducted calculations led to identical results for non-orthogonal and orthogonal operator sets. Spin completeness was found to be of great importance. Spin-incomplete cluster operators led to significant errors in both the correlation energies and the FCI overlap.

ARTICLE HISTORY

Received 6 October 2021
Accepted 7 November 2021

KEYWORDS

Coupled cluster; configuration interaction; spin adaptation; spin completeness; open shell
1. Introduction

Since its development in the late 1950s [1, 2], coupled cluster (CC) has emerged as one of the most accurate quantum chemical methods with applications in all aspects of modern quantum chemistry.

For more than 25 years, it is well understood that spin contamination appears in open-shell CC wave functions (even for pure-spin references, e.g. ROHF) [3–5] and leads to a broken spin symmetry. Ever since, research efforts were applied to solve the latter problem [6–41]. Several advancements were made in the special case of a closed-shell reference [42–55]. This work, however, shall focus on the more general open-shell case, only.

Standard spin-orbital implementations of open-shell CC are readily available in the literature (e.g. Mukherjee et al. [6]) at the cost of spin contamination. To remove the latter, spin-projected [7, 8] and spin-restricted [9–16] methods were developed. Neogrády et al. derived spin-adapted variants of linear CCSD [9], full CCSD [10] and CCSD(T) employing a perturbative triples correction [11] for the doublet spin state while a spin-restricted augmentation of the standard spin-orbital approach was developed by Szalay and Gauß [12]. In their approach, spin constraints, which guarantee a correct spin expectation value, are forced on the CC wave function. To make the approach computationally feasible, the latter constraints are only followed in a truncated subspace such that the approach is not rigorously spin-adapted. Later, a comparison between spin-restricted and rigorously spin-adapted approaches was given [15], the approach was augmented to treat excited states [13], include full triples [14] and has been used in an explicitly correlated R12 framework [16].

Furthermore, Knowles and Werner developed partially spin-adapted CCSD variants [17, 18], where a mix of spin-orbital and spatial substitution operators was used such that the linear CC terms were rigorously spin-adapted. The non-linear terms, however, were not.

In contrast to spin-orbital approaches, the idea of a symmetry-adapted or spin-adapted cluster operator $\hat{T}$, originally suggested by Nakatsuji and Hirao [19], is to define a cluster operator, which commutes with both the total spin operator $\hat{S}^2$ and the spin projection $\hat{S}_z$ with
\[
\left[ \hat{T}, \hat{S}^2 \right] = \left[ \hat{T}, \hat{S}_z \right] = 0, \tag{1}
\]
such that spin quantum numbers $S$ and $S_z$ are automatically conserved. Published approaches may be sorted into the following groups:

(i) plain non-orthogonal approaches [20] employing solitary spatial substitutions $\hat{E}$, which lead to non-orthogonal CSFs,

(ii) orthogonal approaches [21–29] using machinery of unitary group theory to derive linear combinations of spatial substitutions $\hat{E}$, which generate orthogonal CSFs,

(iii) explicitly normal-ordered approaches [30–32], where the wave operator $\hat{e}^\dagger$ is assumed to be normal-ordered, as originally suggested in the late 1970s by Lindgren [56] and

(iv) explicitly normal-ordered approaches reintroducing contractions of spectating substitutions via a combinatoric cluster expansion [33–36], which has recently been augmented to determine first order properties and hyperfine coupling constants [37–39].

All of the aforementioned implementations seem to be limited to the doubles truncation and $S \leq 1$ (triplet spin state). This is probably due to

1. the need for spectating substitutions to reach spin completeness, which complicates the CC theory itself (non-vanishing $\multimap$ and $\multimap$ contractions lead to a non-terminating BCH series) and

2. the need to reformulate the cluster operator for every spin quantum number $S$ as well as its increasing linear dependence with increasing substitution rank.

In our previous work [40], we proposed a solution to (2) based on Löwdin’s [57, 58] projection operator method of spin eigenfunction generation. We developed a canonical operator generation scheme to readily build non-orthogonal cluster operators of arbitrary order, which are applicable to arbitrary (high spin) states. The generated cluster operators, denoted Löwdin-based operators, however, are in no way unique. Due to the increasing linear dependency of the spatial substitution operators with increasing substitution rank, very different sets of same-space-spanning cluster operators are obtainable. Another source of non-uniqueness is an optional subsequent orthogonalisation of the obtained operator.

It is the intent of this work to systematically investigate the effects of spin-adapted cluster operator differences on the wave function quality and the amount of recovered correlation within the non-orthogonal as well as the orthogonal spin-adapted open-shell CC framework. Special attention will be given to our automatically generated Löwdin-based cluster operators [40].

2. Methodology

In this work, several differing sets of spatial substitution operators are applied in CI and CC calculations. The
results are analysed with respect to (i) the wave function quality through means of the overlap to the full CI (FCI) wave function and (ii) the amount of recovered correlation energy compared to FCI. The analysed operator sets are composed of different amounts of spectators as well as orthogonal or non-orthogonal operators. In the following sections, we discuss the generation of various different spin-adapted operator sets.

### 2.1. Generating operator sets

Throughout this paper, we will denote doubly occupied spatial orbitals as \( i, j, k \ldots \in \mathcal{O} \), singly occupied spatial orbitals as \( v, w, x \ldots \in \mathcal{A} \) and empty virtual spatial orbitals as \( \alpha, \beta, \gamma \ldots \in \mathcal{V} \). Arbitrary spatial substitutions \( \hat{E} \) may incorporate substitutions from the joined spaces \( p_1, p_2, p_3 \ldots \in \mathcal{O} \cup \mathcal{A} \) to the joined spaces \( q_1, q_2, q_3 \ldots \in \mathcal{A} \cup \mathcal{V} \). They may be defined by the spin integration of spin-orbital substitution operators \( \hat{X} \) via

\[
\hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{p_1 \ldots p_v} = \sum_{\sigma_1, \sigma_2, \ldots} \sum_{\sigma_v} \hat{X}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{p_1 \sigma_1 \ldots p_v \sigma_v}
\]

\[
= \sum_{\sigma_1, \sigma_2, \ldots} \hat{a}^\dagger_{\sigma_q} \left( \left( \sum_{\sigma_v} \hat{a}_{\sigma_p} \right) \ldots \right) \hat{a}_{\sigma_q}
\tag{2}
\]

Since (2) employs pairwise spin summations only, it is

\[
\hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{p_1 \ldots p_v} = \hat{E}^{(\mathcal{Q}, \mathcal{V})}_{p_1 \ldots p_v} \quad \forall p = \hat{Q} \in \mathcal{S}_v,
\tag{3}
\]

where \( \hat{P} \) and \( \hat{Q} \) denote arbitrary index permutations of the symmetric group \( \mathcal{S}_v \). Therefore, all non-trivial identical spatial substitution operators of rank \( v \) are e.g. given by

\[
\bigcup_{\hat{P} \in \mathcal{S}_v} \hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{\hat{P} p_1 \ldots p_v},
\tag{4}
\]

where all annihilators are sorted in ascending order.

In this work, a total of three different (but complete) operator sets employing solitary spatial substitutions \( \hat{E} \) were generated. For the first two sets, we generated all substitutions \( \hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{\hat{P} p_1 \ldots p_v} \) for all \( \hat{P} \in \mathcal{S}_v \) and for a \( v \) of 1 up to the number of electrons \( n \). These substitutions were then sorted into subsets of configuration generating operators

\[
\mathcal{E}_\lambda = \left\{ \hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{\ldots} \mid \Lambda \left( \hat{E}^{(\mathcal{O}, \mathcal{A}, \mathcal{V})}_{\ldots} \right) = |\lambda\rangle \right\}
\tag{5}
\]

with \( \Lambda(|\Phi\rangle) \) extracting the spatial occupation (configuration) of determinant \( |\Phi\rangle \) and \( |\Psi_0\rangle \) the reference CSF. These subsets are usually much larger than the required spin degeneracy of \( f(O, S) \) for the given spatial configuration (see e.g. [59]) with

\[
f(O, S) = \left( \frac{O}{2} - S \right) - \left( \frac{O}{2} - S - 1 \right)
\tag{6}
\]

for spin quantum number \( S \) and \( O \) open shells. This way we may choose different operators from the subsets to constitute our cluster operators while spanning the same space. In order to build two operator sets, which are as different as possible, we sort the individual \( \hat{E}_\lambda \) with respect to operator rank ascending and descending. From the sorted subsets, we may then choose (i) the first \( f(O, S) \) linearly independent operators such that we arrive at a final set incorporating a minimal amount of spectator indices – denoted by \( \hat{E}_{\text{min}} \) – or (ii) the last \( f(O, S) \) linearly independent operators such that we arrive at a final set incorporating a maximal amount of spectators – denoted by \( \hat{E}_{\text{max}} \).

The third operator set is composed of Löwdin-based operators, which were generated using an automated approach we explicitly derived in our previous work [40]. In this approach, a set of distinct \( \hat{E} \) operator prototypes is permuted and/or augmented utilising Löwdin’s projection operator method for the generation of non-orthogonal spin eigenfunctions [57, 58] to arrive at a linear independent and spin-complete cluster operator. In the following, we will refer to this operator set as \( \hat{E}_{\text{Löw}} \).

### 2.2. (Non-) orthogonality

Since the generated operator sets are composed of single spatial substitutions \( \hat{E} \), they produce non-orthogonal CSFs in general. In the literature, several ansätze utilising orthogonal linear combinations of \( \hat{E} \) substitution operators are known (see e.g. [21, 28]). In terms of complexity and efficiency, non-orthogonal cluster operators are more convenient than orthogonalised cluster operators [54, 60] as may be easily recognised by their simpler form.

For a linear wavefunction ansatz like, e.g. CI the wavefunction and energy are directly governed by the spanned function space. For a non-linear wavefunction approach like, e.g. CC this relation might a priori not be completely obvious. To this end, we take a closer look and consider a non-orthogonal spin-adapted cluster operator \( \hat{T} \) with

\[
\hat{T} = \mathbf{t}^\top \mathbf{E} = \sum_i t_i \hat{E}_i
\tag{7}
\]

where \( \mathbf{t} \) and \( \mathbf{E} \) shall denote vectors composed of the amplitudes \( t_i \) and the non-orthogonal spatial substitution operators \( \hat{E}_i \), respectively. The same space spanning
orthogonal operators $\mathbf{E}'$ may be assembled by

$$\mathbf{E}' = \mathbf{X} \mathbf{E}. \quad (8)$$

Assuming the $\{\hat{E}_i\}$ to be linearly independent also the inverse transformation $\mathbf{X}^{-1}$ exists. This allows for the insertion of the identity $1 = \mathbf{X}^{-1} \mathbf{X}$ in (7) and leads to

$$\hat{T} = t^T \mathbf{X}^{-1} \mathbf{X} \mathbf{E} = t^T \mathbf{E}' = \hat{T}', \quad (9)$$

such that the change of operator basis $\mathbf{E} \rightarrow \mathbf{E}'$ is accompanied by a change of its representation $t \rightarrow t'$. For the same arguments, the powers of $\hat{T}$ and $\hat{T}'$ remain identical. Therefore, any CC calculation (if properly converged) involving either $\hat{T}$ or $\hat{T}'$ should lead to identical wave functions.

In case of the three operator sets $\hat{E}_\text{Low}$, $\hat{E}_\text{min}$ and $\hat{E}_\text{max}$ (c.f. Section 2.1), identical results are expected for their orthogonalised counterparts $\hat{E}'_\text{Low}$, $\hat{E}'_\text{min}$ and $\hat{E}'_\text{max}$ respectively. However, it is important to note that different results are (as in the non-orthogonal case) expected for a direct comparison of $\hat{E}'_\text{Low}$, $\hat{E}'_\text{min}$ and $\hat{E}'_\text{max}$ since they contain different substitutions which can not be transformed into one another by a linear transformation. To summarise, it is $\hat{E}_M \sim \hat{E}_M'$ while $\hat{E}_M \not\sim \hat{E}_N$ and $\hat{E}_M' \not\sim \hat{E}_N$ for $M \neq N$.

To check these equalities (and inequalities), all generated operator sets were explicitly orthonormalised using an appropriately chosen transformation $\mathbf{X}$ which transforms to an orthonormal CSF basis $\mathbf{X} \mathbf{E} |\Psi_0\rangle$. This basis was constructed by merging spatial configurations with the appropriate number of spin eigenfunctions generated via a genealogical coupling algorithm (see e.g. [59]). CC calculations employing both, the non-orthogonal as well as the orthogonalised cluster operator were then compared as outlined in the following sections.

### 2.3. CI and CC in FCI representation

In a general complete non-orthogonal basis set $\{|i\rangle\}$, arbitrary operators $\hat{O}$ are given by

$$\hat{O} = \sum_{ij} |i\rangle \left( \mathbf{S}^{-1} \cdot \mathbf{O} \cdot \mathbf{S}^{-1} \right)_{ij} \langle j|, \quad (10)$$

where $\mathbf{S}^{-1}$ denotes the inverse overlap matrix $\mathbf{S}$ while the elements of $\mathbf{O}$ are given by

$$\left( \mathbf{O} \right)_{ij} = \langle i|\hat{O}|j\rangle. \quad (11)$$

Therefore, representation of general operator products $\hat{O}_1 \cdot \hat{O}_2 \cdots \hat{O}_m$ are given by

$$\langle i|\hat{O}_1 \hat{O}_2 \cdots \hat{O}_m|j\rangle = \left( \mathbf{O}_1 \cdot \mathbf{S}^{-1} \cdot \mathbf{O}_2 \cdot \mathbf{S}^{-1} \cdots \mathbf{S}^{-1} \mathbf{O}_m \right)_{ij}. \quad (12)$$

Please note that an orthonormal CSF basis leaves $\mathbf{S} = \mathbf{S}^{-1} = \mathbf{1}$ such that (12) becomes the much more familiar form of matrix representation products $\mathbf{O}_1 \cdot \mathbf{O}_2 \cdots \cdot \mathbf{O}_m$ only.

In general, the CI and CC wave functions are given by

$$|\Psi_\text{CI}\rangle = (1 + \hat{C}) |\Psi_0\rangle \quad (13)$$

$$|\Psi_\text{CC}\rangle = e^{\hat{T}} |\Psi_0\rangle, \quad (14)$$

respectively, where $|\Psi_0\rangle$ denotes the reference CSF as well as $\hat{C}$ and $\hat{T}$ the CI and cluster operators, respectively. The latter are composed of the particular $\hat{E}$ operators from Section 2.1.

All calculations (CI and CC) analysed in this paper were conducted in the complete FCI CSF basis such that the cluster operator $\hat{T}$, the CI operator $\hat{C}$, the Hamiltonian $\hat{H}$ as well as the identity 1 were represented in this basis as matrix representations $\hat{T}$, $\hat{C}$, $\hat{H}$ and $\mathbf{S}$, respectively.

Represented in a general non-orthogonal basis, the CI equations take the form

$$\left( \hat{H} + \mathbf{H} \cdot \mathbf{S}^{-1} \cdot \mathbf{C} \right)_{X,0} - E_{\text{CI}} \left( \mathbf{S} + \mathbf{C} \right)_{X,0} = 0, \quad (15)$$

with $X \in \mathbb{N}_0$, $X \leq |\text{ls}(\hat{E})| |\Psi_0\rangle|$. The reference CSF $|\Psi_0\rangle$ associated with the energy projection is chosen to be the first basis function $X = 0$. Later basis functions $X > 0$ correspond to the residual projections.

Analogously, the CC equations are given by the BCH series now involving matrix commutators only with

$$\left( \hat{H} + \mathbf{H} \cdot \mathbf{S}^{-1} \cdot \mathbf{T} - \mathbf{T} \cdot \mathbf{S}^{-1} \cdot \mathbf{H} + \cdots \right)_{X,0} = \delta_{X,0} E_{\text{CC}} \quad (16)$$

In this work, the linear (CI) as well as non-linear (CC) equation systems (15) and (16) were solved using the GNU Scientific Library [61] and its gsl_multiroot_fsolver_hybrids module.

### 2.4. Analytics: spin projection error

To check if the used operator sets are producing true spin eigenfunctions, an error estimate $\epsilon$ with

$$\epsilon = \sqrt{1 - \langle \Psi_{\text{CI/CC}}^\dagger \mathbf{S} \Psi_{\text{CI/CC}} \rangle} \quad (17)$$
analogously to [62] was calculated. Here, $|\Psi^{S_S}_{\text{CI/CC}}\rangle$ denotes the projected (after normalisation $\langle \Psi_{\text{CI/CC}} |$) wave function with

$$
|\Psi^{S_S}_{\text{CI/CC}}\rangle = \hat{P}^{S_S}_{\text{CSF}} |\Psi_{\text{CI/CC}}\rangle,
$$

where $\hat{P}^{S_S}_{\text{CSF}}$ shall denote a projector onto the full CSF basis of the desired quantum numbers $S$ and $S_z$. As stated in Ref. [62], the spin projection error $\epsilon$ is a much stronger check for spin state purity than e.g. the deviation of the $S^2$ expectation value since the latter might benefit from error cancellations.

### 2.5. Analytics: FCI overlaps

To investigate the influence of different cluster and CI operators on the overall quality of the CI and CC wave functions, the overlap of the latter wave functions to the exact FCI wave function was calculated. Therefore, both wave functions were normalised. In the untruncated CI and CC limits, the wave functions should be identical such that the (absolute) overlap should be exactly 1.

For cases with degenerate FCI ground state solutions, it may not be enough to consider the overlap to only one of the solutions due to arbitrary rotations in the eigenspace. Therefore, following appendix, the overlap was maximised considering all degenerate FCI solutions using the method of Lagrange multipliers.

### 3. Results

In this section, the results of the conducted calculations are summarised and discussed. In particular, FCI overlaps $\langle \hat{\Psi} | \Psi_{\text{FCI}} \rangle$ are visualised using pW values with

$$
pW = - \log_{10}
\left|
1 - |\langle \hat{\Psi} | \Psi_{\text{FCI}} \rangle|
\right|
$$

which approaches infinity for $|\hat{\Psi}| = |\Psi_{\text{FCI}}|$. Similarly, the amount of recovered correlation energy is visualised by pE values given by

$$
pE = - \log_{10}
\left|
1 - \frac{\hat{E}_{\text{corr}}}{\hat{E}_{\text{FCI}}^{\text{corr}}}
\right|
$$

such that a pE value of 2 represents 99%, a value of 3 99.9%, etc. recovered correlation energy.

A series of CI and CC calculations was carried out on top of converged ROHF calculations using the PySCF program package [63, 64] to analyse the wave function quality as well as the amount of recovered correlation energy comparing (A) the different but same-space-spanning non-orthogonal operator sets $\hat{E}_M$ (B) the non-orthogonal ($\hat{E}_M'$) and orthogonal ($\hat{E}_M''$) operator sets and (C) a potentially spin complete, non-orthogonal, spin-adapted operator sets $\hat{E}_M$ for all generated operator sets $M \in \{\text{Low}, \text{max} \}$ (c.f. Section 2.1). Detailed discussions of (A) – (C) are given in the corresponding Sections 3.1–3.3.

Since both CI and CC calculations were conducted in the FCI CSF basis (c.f. Section 2.3), the presented calculations are limited to only the smallest one-particle basis sets. For all the calculations presented throughout Sections 3.1–3.3, the 6-31G basis set [65] was chosen. While we are aware that 6-31G fails to recover significant portions of the correlation of at least some of the investigated atomic systems, we do not expect our discussed findings and conclusions to be invalidated by this. This is because the present study focuses on differences in the span and structure of the many-particle basis (through means of different $\hat{T}$ and $\hat{C}$ operators) and compares results for identically flawed one-particle basis sets. Nevertheless, we can not preclude that some presented observations might be subject to change when larger one-particle basis sets are used.

Due to the overlapping creator and annihilator space of the employed spatial substitutions, the BCH series is not naturally vanishing for quintuply and higher nested commutators. Therefore, all conducted CC calculations were explicitly checked for their correlation energy convergence with increasing BCH series truncation. This truncation was applied such that the correlation energy was consistent within the applied residual mean square threshold. For all calculations except $^3P$ carbon and $^4S$ nitrogen, a BCH series truncation at quadruply nested commutators was sufficient.

### 3.1. Spin-complete non-orthogonal operator sets: spectator influence

To analyse the influence of the different but same-space-spanning sets of spatial substitution operators $\hat{E}_{\text{Low}}, \hat{E}_{\text{min}}$ and $\hat{E}_{\text{max}}$ (c.f. Section 2.1) on the wave function quality as well as the amount of recovered correlation energy, CI and CC calculations of all possible high spin states of the atomic systems Li, Be, B, C and N in the basis set 6-31G [65] were conducted. Results for all of those systems can be found in the supporting information.

For the sake of this section, we shall focus on the two largest systems (C and N), only. In Figures 1 and 3, the pE and pW values for $^3P$, $^5S$ and $^7P$ atomic carbon as well
as $^4S$ and $^6P$ nitrogen are displayed, respectively. Please note that states $^7P$ and $^6P$ were found here (instead of the expected $^7S$ and $^6S$ states) due to the rather small one particle basis 6-31G. In all Figures 1(a–f) and 3(a–d), two horizontal bar diagrams show the magnitude of the pW (right bar diagram) and the pE values (left bar diagram). Furthermore, subfigures on the left-hand side display CI results (indicated by the striped pattern) while subfigures on the right-hand side display CC results.

The results for the spin-adapted and spin-complete operator sets $\hat{E}_{\text{Low}}$, $\hat{E}_{\text{min}}$ and $\hat{E}_{\text{max}}$ are displayed in blue, red and brown, respectively.

In all calculations, the pW values are larger than the corresponding pE values. Due to the quadratic convergence of the correlation energy, pE values were found corresponding pW values. Due to the quadratic convergence of pE and pW toward the FCI limit but did never lead to a slower convergence than observed with CI. For the CI and CC calculations of $^3P$ carbon shown in Figures 1(a,b), respectively, this is clearly recognisable. CI pW values converge with 1.81, 3.88, 4.57, 8.47 and 9.78 units towards the FCI limit while CC($\hat{E}_{\text{max}}$) pW values converge only slightly faster with 1.81, 3.95, 5.16, 8.83 and 10.80 units to the FCI limit when compared to e.g. CC($\hat{E}_{\text{Low}}$) with 1.81, 3.99, 6.72, 10.13 and 13.00 units.

The reason for this unfavourable behaviour of the SASC-CC($\hat{E}_{\text{max}}$) calculations can be found in the non-linear product terms $\hat{T}^2$, $\hat{T}^3$, etc. For an increased amount of spectators, an increased amount of these product terms vanishes. Ultimately, for the worst case $T^n = 0 \forall n \geq 1$ this leads to a completely linear CC wave function ansatz analogous to CI. Consider e.g. the spatial configuration $|1\rangle|2\rangle|3\rangle$ composed of spatial orbitals 1, 2 and 3 reachable from the doublet reference CSF $|\alpha\rangle$ via a double substitution. Here we denote alpha and beta spin states via non-overlined and overlined indices. The spin degeneracy for 3 open shells and $S = S_z = \frac{1}{2}$ requires

$$f\left(3, \frac{1}{2}\right) = \left(\begin{array}{c}3 \\ 1 \\ 0 \end{array}\right) - \left(\begin{array}{c}3 \\ 0 \\ 0 \end{array}\right) = 2$$

operators to span the complete spin space. In total, the three operators

$$\hat{E}_{\text{00}}^{23} |001\rangle = |123\rangle - |1\bar{2}3\rangle$$
$$\hat{E}_{\text{001}}^{12} |001\rangle = |123\rangle - |1\bar{2}3\rangle$$
$$\hat{E}_{\text{001}}^{13} |001\rangle = |1\bar{2}3\rangle - |1\bar{2}3\rangle$$

appear in the operator sets $\hat{E}_{\text{Low}}, \hat{E}_{\text{min}}$ and $\hat{E}_{\text{max}}$ as outlined in Figure 2.

The operator sets $\hat{E}_{\text{Low}}$ and $\hat{E}_{\text{min}}$ are composed of the minimal spatial substitution $\hat{E}_{\text{00}}^{23}$ and the second spatial substitution $\hat{E}_{\text{001}}^{12}$ or $\hat{E}_{\text{001}}^{13}$ incorporating the spectating index 1. In contrast, the operator set $\hat{E}_{\text{max}}$ does not include the minimal spatial substitution $\hat{E}_{\text{00}}^{23}$ but only the two spectators. Any product with a single substitution $\hat{E}_{a}^{a}$ from the right, leads to

$$\hat{E}_{00}^{23} \hat{E}_{1}^{a} |0\bar{01}\rangle = (\delta_{a3} - 1) |23a\rangle + (1 - \delta_{a2}) |2\bar{3}a\rangle$$
$$\hat{E}_{001}^{12} \hat{E}_{1}^{a} |0\bar{01}\rangle = 0$$
$$\hat{E}_{001}^{13} \hat{E}_{1}^{a} |0\bar{01}\rangle = 0$$

such that products (…)$\hat{E}_{a}^{a}$ completely vanish for the operator set $\hat{E}_{\text{max}}$. In case of the operator sets $\hat{E}_{\text{Low}}$ and $\hat{E}_{\text{min}}$ this is not entirely the case since the product $\hat{E}_{00}^{23} \hat{E}_{1}^{a}$ does not vanish for arbitrary indices $a \in \mathbb{V}$.

### 3.2. Invariance w.r.t. orthogonalisation for spin complete operator sets

As an additional validation of the present implementation we checked for invariance of the results w.r.t.
Figure 1. Summarised $p_E$ and $p_W$ values for spin-adapted and spin-complete CI (left) and CC (right) calculations for atomic $^3P$, $^5S$, and $^7P$ carbon in the 6-31G* basis set. (a) $^3P$ carbon (CI), (b) $^3P$ carbon (CC), (c) $^5S$ carbon (CI), (d) $^5S$ carbon (CC), (e) $^7P$ carbon (CI), (f) $^7P$ carbon (CC).

the orthogonality of the spin operators for CI and CC approaches. All spin-complete orthogonal CI/CC results are identical to the spin-complete non-orthogonal CI/CC results. Therefore, an analogous discussion to the influence of different amounts of spectators in Section 3.1 may be applied to the orthogonalised operator sets $\hat{E}_{\text{Loew}}$, $\hat{E}_{\text{min}}$, and $\hat{E}_{\text{max}}$.

Results of all conducted CI as well as CC calculations using the orthonormalised operator sets $\hat{E}_{\text{Loew}}$, $\hat{E}_{\text{min}}$, and $\hat{E}_{\text{max}}$ can be found in the supporting information.
3.3. Analysing the influence of spin completeness

As explained in our previous work [40], $\hat{E}$ operator spectators are necessary to reach spin-complete operator sets for open-shell spin-adapted CC calculations. Without the minimal required amount of spectators, a significant spin incompleteness error can arise. To systematically investigate this spin incompleteness effect, we defined three different cluster/CI operator truncation sublevels $(0), (1)$ and $(2)$. Each truncation sublevel may be applied to each operator set $\hat{E}_M$ for each spatial substitution order $\nu$. Here,

$(0)$ denotes the unaltered fully spin-adapted and spin-complete operator set $\hat{E}_M^{(0)} = \hat{E}_M$,

$(1)$ denotes a spin-adapted but possibly spin-incomplete operator set $\hat{E}_M^{(1)}$, where only spectators of nominal tensor rank $\mu$ with $\mu \leq \nu$ are included and

$(2)$ denotes a spin-adapted but spin-incomplete operator set including no spectators at all.

In Table 1, the affiliation of different single and double spatial substitution operators from the operator set $\hat{E}_M$ to the sublevels $(0)$, $(1)$ and $(2)$ in the spatial singles $(S)$ as well as singles and doubles $(SD)$ truncation is shown.

For the SD truncation, sublevel $(0)$ includes all spectators – also operators of tensor ranks 3 and 4 for doubles – sublevel $(1)$ readily stops at tensor rank 2 operators but includes all rank 2 spectators (of spatial substitution order $\nu = 1$) and sublevel $(2)$ includes no spectators at all.

Figure 2. Affiliation diagram of the three operators $\hat{E}_{00}^{23}, \hat{E}_{001}^{123}$ and $\hat{E}_{001}^{132}$ to the $\hat{E}_{\text{Löw}}, \hat{E}_{\text{min}}$ and $\hat{E}_{\text{max}}$ operator sets.

Figure 3. Summarised $pE$ and $pW$ values for spin-adapted and spin-complete CI (left) and CC (right) calculations for atomic $^4$S and $^6$P nitrogen in the 6-31G [65] basis set. (a) $^4$S nitrogen (CI), (b) $^4$S nitrogen (CC), (c) $^6$P nitrogen (CI), (d) $^6$P nitrogen (CC).
the FCI limit), as spin projection error of zero is obtained. For an untruncated cluster operator (SDTQ5), calculated according to equations (19) and (20). Results of recovered correlation energy, pW and pE values were within the precision limit proving their spin state purity. In contrast, spin-orbital CC calculations lead to spin projection errors in the order of $10^{-2}$ to $10^{-8}$ for truncation S to SDTQ5Q. For an untruncated cluster operator (SDTQ5Q6, the FCI limit), a spin projection error of zero is obtained.

To investigate the effect of the different sublevels (0), (1) and (2) on the wave function as well as the amount of recovered correlation energy, pW and pE values were calculated according to equations (19) and (20). Results of all conducted calculations (Li to N) can be found in the supporting information.

In Figure 4, the obtained results for the $^4$S and $^6$P high spin states of atomic nitrogen are shown. The shown results were obtained from orthogonal (left) and non-orthogonal (right) CC calculations employing the operator set $\hat{E}_{\text{Low}}$. As in the previous sections, bar charts advancing to the right show pW values while bar charts advancing to the left show pE values. To distinguish among the three truncation sublevels (0), (1) and (2), blue colours with decreasing opacity were used. Here, dark blue resembles truncation sublevel (0), light blue sublevel (1) and white sublevel (2).

Clearly, truncation sublevel (2) leads to very inaccurate results compared to sublevels (0) and (1). For the shown example (neglecting singles), pE values are found between 1.06–8.32 units smaller than sublevels (0) and (1) while pW values are found between 0.64–10.76 units smaller than sublevels (0) and (1). The comparison of sublevels (0) or (1) and sublevel (2), is however not on the same footing since even for a full spatial substitution rank $\nu = n$, sublevel (2) can never reach the FCI limit while sublevels (0) and (1) readily converge to the same FCI solution.

Comparing sublevels (0) and (1), only minor differences are recognisable in the non-orthogonal case. There are however certain exceptions like, e.g. $^4$Snitrogen CCSD. Here, pE values of 2.20 and 1.84 are obtained, which indicate 99.37% and 98.42% of recovered correlation energy (w.r.t. FCI), respectively. This error of roughly 1% in the correlation energy might be of importance when approaching chemical accuracy. In most of the conducted non-orthogonal calculations, however, the difference between sublevels (0) and (1) seems to be negligible. This effect is of particular interest since the high rank spectators in sublevel (0), which are neglected in sublevel (1), are expected to be the computationally most expensive terms of the CC equations. Since only small atomic calculations were conducted in this work, this effect might not necessarily carry on to larger molecular systems. Here, the difference between these sublevels might be of greater importance. This will be subject to further investigations.

In the orthogonal case, the results of sublevel (0) (as discussed in Section 3.2) are identical to the non-orthogonal results. For sublevels (1) and (2), spectators of different tensor ranks are neglected. In the orthogonal cluster operator, which is composed of linear combinations of spatial substitutions, all orthogonal substitutions including at least one neglected spectator are completely ignored. Therefore, sublevels (1) and (2) may lead to fewer amplitudes in the orthogonal case compared to the non-orthogonal case resulting in smaller pE and pW

### Table 1. Truncation sublevel affiliation of different spatial substitution operator prototypes of spatial substitution order 1 and 2.

| Order $\nu$ | Rank | Operators | Sublevel (0) | Sublevel (1) | Sublevel (2) |
|-------------|------|-----------|--------------|--------------|--------------|
| 1           | 2    | $\hat{E}_{\nu}^{ii}$ | ✓            | ✓            | ✓            |
| 2           | 3    | $\hat{E}_{\nu}^{ijv}$ | ✓            | ✓            | ✓            |

### Table 2. Summarised spin projection errors for spin-orbital and spin-adapted non-orthogonal CC calculations employing the cluster operator truncation sublevels (0), (1) and (2) for the operator set $\hat{E}_{\text{Low}}$ of atomic $^3$P carbon in the 6-31G(65) basis set.

| Truncation | Spin-orbital CC | Spin-Adapted CC |
|------------|-----------------|-----------------|
| S          | $2.07 \cdot 10^{-2}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
| SD         | $1.79 \cdot 10^{-3}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
| SDT        | $8.26 \cdot 10^{-5}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
| SDTQ       | $1.75 \cdot 10^{-6}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
| SDTQ5      | $1.96 \cdot 10^{-8}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
| SDTQ56     | $3.46 \cdot 10^{-16}$ | $< 10^{-17}$ | $< 10^{-17}$ | $< 10^{-17}$ |
values in the orthogonal sublevels (1) and (2) compared to the non-orthogonal case.

4. Conclusion

In this work, atomic test calculations of spin-adapted CI and CC were investigated in terms of

(a) the wave function quality (through means of the overlap to the FCI wave function) and
(b) the amount of recovered correlation energy (w.r.t. FCI).

Both (a) and (b) were compared for CI and CC calculations using different sets of spatial substitution operators composed of different amounts of spectators as well as orthogonal and non-orthogonal substitutions.

In Section 3.1, the influence of different amounts of spectators in spatial substitution operator sets spanning identical linear spaces was investigated. In total, three different but same-space-spanning operator sets \( \hat{E}_{\text{Low}}, \hat{E}_{\text{min}} \) and \( \hat{E}_{\text{max}} \) were compared. While \( \hat{E}_{\text{min}} \) and \( \hat{E}_{\text{max}} \) were composed of a minimal and maximal amount of spectators, the set \( \hat{E}_{\text{Low}} \) was generated by an automated operator generation approach explicitly discussed in our previous work \([40]\). As expected, all CI calculations show identical results for all operator sets, while CC (due to its non-linear wave function ansatz) showed differences in both the FCI overlap and the amount of recovered correlation energy. In the present study, the best results were found for the operator sets \( \hat{E}_{\text{Low}} \) and \( \hat{E}_{\text{min}} \), which were almost identical in most scenarios. Operator set \( \hat{E}_{\text{max}} \) led to significantly flawed CC results with errors in the correlation energy of up to 0.32% (w.r.t. FCI).

In Section 3.2, the CC results for sets of orthogonal and non-orthogonal spatial substitution operators were compared. As explicitly discussed in Section 2.2, cluster operators linked by an invertible linear transformation should
lead to identical results. As expected, identical results for spin-complete orthogonal and non-orthogonal CC were obtained. Future open-shell spin-adapted CC developments may therefore benefit from the usage of the simpler non-orthogonal operators [40] over the more complicated orthogonalised form without any accuracy gain or loss.

Finally, in Section 3.3, the effect of spin incompleteness on the correlation energy and the wave function quality was investigated. Three truncation sublevels (0), (1) and (2) for an increasingly insufficient spin space were defined for each spatial substitution truncation level. Here, sublevel (0) denotes the fully spin-complete cluster operator, sublevel (1) neglects operators with nominal tensor ranks higher than the applied spatial truncation while sublevel (2) denotes a strongly spin-incomplete cluster operator including no spectators at all. The results for sublevels (0) and (1) were found to be qualitatively identical for non-orthogonal cluster operators with quantitative agreement in most cases. The results for sublevel (2), however, showed large errors of several orders of magnitude in both the recovered correlation energy and the wave function emphasising the importance of spin completeness in CI and CC calculations.

**Note**

1. Please note that rank here implies the nominal tensor rank and not the order of spatial substitution since the latter would be identical for all operators in $\hat{E}_\lambda$.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by Evangelisches Studienwerk Viligist.

**ORCID**

Nils Herrmann http://orcid.org/0000-0002-6468-9180
Michael Hanrath http://orcid.org/0000-0003-3188-3961

**References**

[1] F. Coester, Nucl. Phys. 7, 421 (1958). doi:10.1016/0029-5582(58)90280-3
[2] F. Coester and H. Kümmel, Nucl. Phys. 9, 225 (1958). doi:10.1016/0029-5582(58)90394-8
[3] J.F. Stanton, J. Chem. Phys. 101, 371 (1994). doi:10.1063/1.468144
[4] A.I. Krylov, J. Chem. Phys. 113, 6052 (2000). doi:10.1063/1.1308557
[5] X. Li and J. Paldus, Int. J. Quantum Chem. 77, 281 (2000). doi:10.1002/(ISSN)1097-461X
[6] D. Mukherjee, R.K. Moitra and A. Mukhopadhyay, Mol. Phys. 30, 1861 (1975). doi:10.1080/00268977500103351
[7] M. Rittby and R.J. Bartlett, J. Phys. Chem. 92, 3033 (1988). doi:10.1021/j100322a004
[8] H. Yuan and D. Cremer, Theor. Chem. Acc. 105, 132 (2000). doi:10.1007/s002100010195
[9] P. Neogrády, M. Urban and I. Hubač, J. Chem. Phys. 97, 5074 (1992). doi:10.1063/1.463828
[10] P. Neogrády, M. Urban and I. Hubač, J. Chem. Phys. 100, 3706 (1994). doi:10.1063/1.466359
[11] P. Neogrády, M. Urban, Int. J. Quantum Chem. 55, 187 (1995). doi:10.1002/qua.5050214
[12] P.G. Szalay, J. Gauss, J. Chem. Phys. 107, 9028 (1997). doi:10.1063/1.475220
[13] P.G. Szalay and J. Gauss, J. Chem. Phys. 112, 4027 (2000). doi:10.1063/1.480952
[14] I. Berente, P.G. Szalay and J. Gauss, J. Chem. Phys. 117, 7872 (2002). doi:10.1063/1.1506916
[15] M. Heckert, O. Heun, J. Gauss and P.G. Szalay, J. Chem. Phys. 124, 124105 (2006). doi:10.1063/1.2179070
[16] J.J. Wilke and H.F. Schaefer III, J. Chem. Theory Comput. 7, 2416 (2011). doi:10.1021/ct200292z
[17] P.J. Knowles, C. Hampel and H.J. Werner, J. Chem. Phys. 99, 5219 (1993). doi:10.1063/1.465990
[18] P.J. Knowles, C. Hampel and H.-J. Werner, J. Chem. Phys. 112, 3106 (2000). doi:10.1063/1.480886
[19] H. Nakatsuji and K. Hirao, J. Chem. Phys. 68, 2053 (1978). doi:10.1063/1.436028
[20] G.L. Janssen and H.F. Schaefer III, Theor. Chim. Acta 79, 1 (1991). doi:10.1007/BF01113327
[21] X. Li and J. Paldus, Int. J. Quantum Chem. 48, 269 (1993). doi:10.1002/issn.1097-461X
[22] X. Li and J. Paldus, J. Chem. Phys. 101, 8812 (1994). doi:10.1063/1.468074
[23] P. Piecuch, X. Li and J. Paldus, Chem. Phys. Lett. 230, 377 (1994). doi:10.1016/0009-2614(94)01167-2
[24] X. Li, P. Piecuch and J. Paldus, Chem. Phys. Lett. 224, 267 (1994). doi:10.1016/0009-2614(94)00563-X
[25] X. Li and J. Paldus, J. Chem. Phys. 102, 8059 (1995). doi:10.1063/1.469005
[26] X. Li and J. Paldus, J. Chem. Phys. 102, 2013 (1995). doi:10.1063/1.468766
[27] X. Li and J. Paldus, J. Chem. Phys. 103, 6536 (1995). doi:10.1063/1.470380
[28] B. Jeziorski, J. Paldus and P. Jankowski, Int. J. Quantum Chem. 56, 129 (1995). doi:10.1002/issn.1097-461X
[29] P. Jankowski and B. Jeziorski, J. Chem. Phys. 111, 1857 (1999). doi:10.1063/1.479511
[30] M. Nooijen and R.J. Bartlett, J. Chem. Phys. 104, 2652 (1996). doi:10.1063/1.471010
[31] M. Nooijen and V. Lotrich, Comput. Theor. Chem. 547, 253 (2001). doi:10.1016/S0166-1280(01)00475-4
[32] S. Sen, A. Shee and D. Mukherjee, J. Chem. Phys. 137, 074104 (2012). doi:10.1063/1.4742058
[33] D. Datta and D. Mukherjee, Int. J. Quantum Chem. 108, 2211 (2008). doi:10.1002/qua.21738
[34] D. Datta and D. Mukherjee, J. Chem. Phys. 131, 044124 (2009). doi:10.1063/1.3185356
[35] D. Datta and D. Mukherjee, J. Chem. Phys. 134, 054122 (2011). doi:10.1063/1.3537740
[36] D. Datta and J. Gauss, J. Chem. Theory Comput. 9, 2639 (2013). doi:10.1021/ct400216h
[37] D. Datta and J. Gauss, J. Chem. Phys. 141, 104102 (2014). doi:10.1063/1.4894773
[38] D. Datta and J. Gauss, J. Chem. Phys. 143, 011101 (2015). doi:10.1063/1.4923436
[39] D. Datta and J. Gauss, J. Chem. Theory Comput. 15, 1572 (2019). doi:10.1021/acs.jctc.8b01048
[40] N. Herrmann and M. Hanrath, J. Chem. Phys. 153, 164114 (2020). doi:10.1063/5.0026762
[41] A.I. Krylov, The Quantum Chemistry of Open-Shell Species (John Wiley & Sons, Ltd., 2017). Chap. 4, pp. 151–224.
[42] J. Paldus, J. Chem. Phys. 67, 303 (1977). doi:10.1063/1.434526
[43] B.G. Adams and J. Paldus, Phys. Rev. A 20, 1 (1979). doi:10.1103/PhysRevA.20.1
[44] S. Ghosh, D. Mukherjee and S.N. Bhattacharyya, Phys. Phys. 72, 161 (1982). doi:10.1063/1.87077-8
[45] S.S. Aadnan, S. Bhattacharyya and D. Mukherjee, Chem. Phys. 85, 204 (1982). doi:10.1006/istc.1992.80333-3
[46] M. Takahashi and J. Paldus, J. Chem. Phys. 85, 1486 (1985). doi:10.1063/1.451241
[47] B. Jeziorski and J. Paldus, J. Chem. Phys. 88, 5673 (1988). doi:10.1063/1.454528
[48] P. Piecuch and J. Paldus, Int. J. Quantum Chem. 36, 429 (1989). doi:10.1002/issn:1097-461X
[49] P. Piecuch and J. Paldus, Theor. Chim. Acta 78, 65 (1990). doi:10.1007/BF01119191
[50] P. Piecuch and J. Paldus, Theor. Chim. Acta 83, 69 (1992). doi:10.1007/BF00113244
[51] P. Piecuch, R. Tobola and J. Paldus, Int. J. Quantum Chem. 55, 133 (1995). doi:10.1002/issn:1097-461X
[52] A.E. Kondo, P. Piecuch and J. Paldus, J. Chem. Phys. 102, 6511 (1995). doi:10.1063/1.469365
[53] A.E. Kondo, P. Piecuch and J. Paldus, J. Chem. Phys. 104, 8566 (1996). doi:10.1063/1.471546
[54] D.A. Matthews, J. Gauss and J.F. Stanton, J. Chem. Theory Comput. 9, 2567 (2013). doi:10.1021/ct301024v
[55] D.A. Matthews and J.F. Stanton, J. Chem. Phys. 142, 064108 (2015). doi:10.1063/1.4907278
[56] I. Lindgren, Int. J. Quantum Chem. 14, 33 (1978).
[57] P.-O. Löwdin, Phys. Rev. 97, 1509 (1955). doi:10.1103/PhysRev.97.1509
[58] P.-O. Löwdin, Rev. Mod. Phys. 36, 966 (1964). doi:10.1103/RevModPhys.36.966
[59] R. Pauncz, Spin Eigenfunctions Construction and Use, 1st ed. (Plenum Press, New York, 1979).
[60] P. Springer, D. Matthews and P. Bientinesi, ACM Trans. Math. Software 45, 1 (2019). doi:10.1145/3301319
[61] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth and F. Rossi, GNU Scientific Library Reference Manual, 3rd ed. (2009).
[62] M. Hanrath and A. Engels-Putzka, Theor. Chem. Acc. 122, 197 (2009). doi:10.1007/s00214-008-0499-3
[63] Q. Sun, T.C. Berkelbach, N.S. Blunt, G.H. Booth, S. Guo, Z. Li, J. Liu, J.D. McClain, E.R. Sayfutyarova, S. Sharma, S. Wouters and G.K.L. Chan, WIREs Comput. Mol. Sci. 8, e1340 (2018). doi:10.1002/wcms.2018.8.issue-1
[64] Q. Sun, J. Comput. Chem. 36, 1664 (2015). doi:10.1002/jcc.v36.22
[65] W.J. Hehre, K. Ditchfield and J.A. Pople, J. Chem. Phys. 56, 2257 (1972). doi:10.1063/1.1677527

Appendix

Appendix. Overlap maximisation

Consider a set \(|\Phi_i\rangle\) for \(i = 1, \ldots, m\) of degenerate eigenfunctions of the Hamiltonian with

\[ \hat{H} |\Phi_i\rangle = E_i |\Phi_i\rangle \quad \forall_i, \quad (A1) \]

for which

\[ \langle \Phi_j | \Phi_i \rangle = \delta_{ij} \quad (A2) \]

holds. Due to the linearity of the Hamiltonian, it is clear that arbitrary linear combinations \(|\Psi\rangle\) with

\[ |\Psi\rangle = \sum_{i=1}^{m} c_i |\Phi_i\rangle \quad (A3) \]

are also eigenfunctions of \(\hat{H}\) with the same eigenvalue \(E\) – hence the rotational freedom of the eigenspace. For a given CC or CI wave function \(|\Psi\rangle\), the overlap to \(|\Psi\rangle\) may be maximised under the constraint of the latter remaining normalised. An appropriate Lagrangian \(L\) takes the form

\[ L(c_1, \ldots, c_m, \lambda) = \sum_{i=1}^{m} c_i \langle \Phi_i | \Phi_i \rangle + \lambda \left( \sum_{i=1}^{m} |c_i|^2 - 1 \right) \quad (A4) \]

Assuming real coefficients \(c_i\), the partial derivatives with respect to \(c_j\) and \(\lambda\) are given by

\[ \frac{\partial L(c_1, \ldots, c_m, \lambda)}{\partial c_j} = \langle \Phi_j | \Phi_j \rangle + 2\lambda c_j \quad \forall_j \in \{1, \ldots, m\} \quad (A5) \]

\[ \frac{\partial L(c_1, \ldots, c_m, \lambda)}{\partial \lambda} = \sum_{i=1}^{m} c_i^2 - 1 \equiv 0, \quad (A6) \]

respectively. From Equation (A5) for \(j = 1 \ldots m\), it is

\[ c_j = -\frac{\langle \Psi | \Phi_j \rangle}{2\lambda} \quad (A7) \]

Insertion into (A6) then leads to

\[ \lambda = \pm \sqrt{\frac{\sum_{i=1}^{m} \langle \Psi | \Phi_i \rangle^2}{4}}, \quad (A8) \]

such that the maximised (minimised) overlap \(S\) is given by

\[ S = \sum_{i=1}^{m} c_i^2 \langle \Psi | \Phi_i \rangle = \frac{\sum_{i=1}^{m} \langle \Psi | \Phi_i \rangle^2}{-2\lambda} \quad (A9) \]

for a negative (positive) choice of \(\lambda\).