Posted on 31 Mar 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164873477.73577108/v1 — This is a preprint. Version of Record available at https://doi.org/10.1002/qua.27045

A Recursive Formulation of One-Electron Coupling Coefficients for Spin-Adapted Configuration Interaction Calculations Featuring Many Unpaired Electrons

Mihkel Ugandi[1] and Michael Römelt[1]

[1]Humboldt-Universität zu Berlin

February 22, 2024

Abstract

This work reports on a novel computational approach to the efficient evaluation of one-electron coupling coefficients as they are required during spin-adapted electronic structure calculations of the configuration interaction type. The presented approach relies on the equivalence of the representation matrix of excitation operators in the basis of configuration state functions and the representation matrix of permutation operators in the basis of genealogical spin eigenfunctions. After the details of this connection are established for every class of one-electron excitation operator, a recursive scheme to evaluate permutation operator representations originally introduced by Yamanouchi and Kotani is recapitulated. On the basis of this scheme we have developed an efficient algorithm that allows the evaluation of all nonredundant coupling coefficients for systems with 20 unpaired electrons and a total spin of S = 0 within only a few hours on a simple Desktop-PC. Furthermore, a full-CI implementation that utilizes the presented approach to one-electron coupling coefficients is shown to perform well in terms of computational timings for CASCI calculations with comparably large active spaces. More importantly, however, this work paves the way to spin-adapted and configuration driven selected configuration interaction calculations with many unpaired electrons.

A Recursive Formulation of One-Electron Coupling Coefficients for Spin-Adapted Configuration Interaction Calculations Featuring Many Unpaired Electrons

Mihkel Ugandi and Michael Roemelt[a)]
Institut für Chemie, Humboldt-Universität zu Berlin, Brook-Taylor-Str. 2, D-12489 Berlin, Germany

(Dated: 16 March 2022)

This work reports on a novel computational approach to the efficient evaluation of one-electron coupling coefficients as they are required during spin-adapted electronic structure calculations of the configuration interaction type. The presented approach relies on the equivalence of the representation matrix of excitation operators in the basis of configuration state functions and the representation matrix of permutation operators in the basis of genealogical spin eigenfunctions. After the details of this connection are established for every class of one-electron excitation operator, a recursive scheme to evaluate permutation operator representations originally introduced by Yamanouchi and Kotani is recapitulated. On the basis of this scheme we have developed an efficient algorithm that allows the evaluation of all nonredundant coupling coefficients for systems with 20 unpaired electrons and a total spin of $S = 0$ within only a few hours on a simple Desktop-PC. Furthermore, a full-CI implementation that utilizes the presented approach to one-electron coupling coefficients is shown to perform well in terms of computational timings for CASCI calculations with comparably large active spaces. More importantly, however, this work paves the way to spin-adapted and configuration driven selected configuration interaction calculations with many unpaired electrons.

Keywords: Electronic Structure Theory, Spin-Adaptation, Configuration Interaction, Coupling Coefficients

I. INTRODUCTION

Full configuration interaction (Full-CI) and selected configuration interaction (SCI) variants are frequently used in the context of multireference (MR) electronic structure theory which plays an important role in multiple branches of chemistry, e.g. photochemistry and transition metal chemistry.[1,2] Depending on the circumstances it is beneficial to solve the corresponding eigenvalue equation either in the basis of Slater determinants (SDs) or spin adapted configuration state functions (CSFs). A formulation in terms of Slater determinants results in relatively simple expressions for the molecular Hamiltonian matrix elements and allows for an efficient implementation in computer programs,[3–6] in particular when the involved routines are run on graphical processing units (GPUs).[7,8] Yet, it is impossible to directly target specific spin states and the wavefunction representation in terms of Slater determinants can become quite inefficient.[9,10] In contrast, a formulation in terms of CSFs comes at the cost of a more complex logic but results in a more compact wavefunction representation, particularly when antiferromagnetically coupled local spins are present, e.g. in polynuclear transition metal clusters. Formulations that use a mixed CSF/SD approach aim to capture the best of two worlds, the seemingly simple logic during matrix element evaluation, rigorous enforcement of spin symmetry and a compact representation.[11,12] Such approaches have been enabled by the development of routines for the fast transformation

a)Electronic mail: michael.roemelt@hu-berlin.de

of expansion vectors between the two representations.[13]

One of the critical aspects of a CSF-based formulation is the efficient evaluation of one-electron coupling coefficients between spin-adapted basis functions. An approach based on explicit representations of the involved spin eigenfunctions suffers from the factorial scaling of the number of primitive spin functions and is therefore impractical for systems with many unpaired electrons.[9] A more elegant alternative is provided by the graphical unitary group approach (GUGA) that was developed by Moshinsky, Paldus and Shavitt and coworkers in the 1970's and has recently been used by Alavi and coworkers to formulate a spin-adapted version of stochastic CI.[14–19] However, within configuration-driven CI calculations the GUGA has to be applied to all pairs of CSFs leading to numerical bottlenecks due to the fast growing number of CSFs with increasing number of unpaired electrons.

This work reports on the theoretical background and implementation of a recursive scheme to efficiently evaluate matrices of one-electron coupling coefficients between all CSFs of a given configuration based on representation matrices of permutation operators. A similar approach to the calculation of such matrices was established by Rettrup[20] which was later used by Werner and coworkers in various contexts.[21,22] However, in their works the coupling coefficient matrices are constructed as products of matrices that correspond to a single orbital index. While this decomposition serves to reduce the number of required permutation classes and the storage requirement, it doubles the amount of matrix operations during CI calculations. In the algorithm presented here, all different classes of coupling coefficient matrices are constructed as a whole thereby slightly increasing the complexity of the logic. Yet, a key feature of the presented algorithm is that only non-zero matrix elements are explicitly calculated thereby making it remarkably efficient in terms of required computer time. Furthermore, we present a Full-CI program that utilizes large coupling coefficient matrices in a direct fashion which avoids any memory bottlenecks entirely.

II. THEORY

A. Configuration Interaction And Coupling Coefficients

Before the aforementioned recursive scheme is described in detail we briefly discuss the identity of one-electron coupling coefficients and their usage in spin-adapted CI. A general CI wavefunction for N electrons takes the form

$$|\Psi_{CI}\rangle = \sum_I C_I \, |\Phi_I\rangle \tag{1}$$

where in the current context the set of N-electron basis functions $\{\Phi_I\}$ consists of spin-adapted CSFs. Since the dimension of the CSF basis is usually too large to directly diagonalize the molecular Hamiltonian matrix $\mathbf{H}$, the CI problem is solved by means of multiroot Davidson or Lanczos algorithms.[23–25] The central quantity one has to calculate for these algorithms is the vector $\sigma = \mathbf{HC}$. After application of the resolution-of-the-identity (RI) in the CSF space the elements of σ can be written as

$$\sigma_I = \sum_J \sum_{pq} h_{pq} A^{IJ}_{pq} C_J + \frac{1}{2} \sum_{JK} \sum_{pqrs} (pq|rs) A^{IK}_{pq} A^{KJ}_{rq} C_J \tag{2}$$

Here, h_{pq} and $(pq|rs)$ denote molecular one- and two-electron integrals, respectively, while the set of $\{C_J\}$ are the CSF expansion coefficients introduced in equation (1). The one-electron coupling coefficients are given by

$$A^{IK}_{pq} = \langle \Phi_I | E^p_q | \Phi_K \rangle \tag{3}$$

where $E_q^p = \hat{a}_{p\alpha}^\dagger \hat{a}_{q\alpha} + \hat{a}_{p\beta}^\dagger \hat{a}_{q\beta}$ is a spin-traced replacement operator in second quantization. Obviously, the two-electron part of equation (2) is the computationally most demanding part of the σ vector evaluation. However, it can brought into the matrix-form[26]

$$\Delta\sigma_I = \frac{1}{2} Tr(\mathbf{A}^I \mathbf{ID}) \tag{4}$$

where $I_{pq,rs} = (pq|rs)$ and $D_{rs,K} = \sum_J A_{rs}^{KJ} C_J$. Importantly, once the one-electron coupling coefficients are known, the σ vector in equation (2) can be efficiently evaluated using modern linear algebra routines.

B. Symmetry of Coupling Coefficients

With increasing number of molecular orbitals and configurations the number of coupling coefficients that have to be evaluated grows rapidly. However, the numerical value of A_{pq}^{IK} is identical for many CSFs which significantly reduces the storage requirements. Most importantly, two coupling coeffcients A_{pq}^{IK} and $A_{pq}^{I'K'}$ have identical values if the position of the two indices p and q relative to the unpaired electrons is the same and the spin coupling in Φ_I and $\Phi_{I'}$ as well as Φ_K and $\Phi_{K'}$ is identical. This symmetry can be illustrated by means of a simple example with the three CSFs

$$|\Phi_1\rangle = |\mathbf{n} = 22211; S = 0\rangle \tag{5a}$$
$$|\Phi_2\rangle = |22121; 0\rangle \tag{5b}$$
$$|\Phi_3\rangle = |12221; 0\rangle \tag{5c}$$

where each CSF is determined by an occupation number vector $\mathbf{n}$ and a total spin S. Note that since there is only one way of constructing a spin eigenfunction with $S = 0$ from $N = 2$ unpaired electrons, no extra index k for the spin eigenfunction has to be introduced in this case. Generally, the number of independent spin eigenfunctions with N unpaired electrons and total spin S is obtained as

$$f_{NS} = \binom{N}{\frac{1}{2}N - S} - \binom{N}{\frac{1}{2}N - S - 1} \tag{6}$$

In the present example, Φ_1 is connected to Φ_2 and Φ_3 through a single excitation of the DOMO→SOMO type via excitation operators E_3^4 and E_1^4, respectively. Without actually calculating the matrix elements it can be seen that

$$\langle\Phi_2|E_3^4|\Phi_1\rangle = \langle\Phi_3|E_1^4|\Phi_1\rangle \tag{7}$$

since the position of the doubly occupied donor orbital in Φ_1 relative to the singly occupied orbitals in Φ_1 is identical. In other words, one may permute doubly occupied orbitals without having an influence on the single-excitation matrix element as long as there are no singly occupied orbitals between the double occupied orbitals and as long as these permutations affect both involved CSFs. Therefore, for a given number of unpaired electrons N and total spin S one only needs to compute and store the coupling coefficients for all possible combinations of $N + 1$ relative donor positions and N relative acceptor positions. This reduces the computation and storage requirement to $(N + 1) \times N$ coupling coefficient matrices of dimension $(f_{NS} \times f_{NS})$ to represent all DOMO→SOMO excitations. Analogous arguments can be made for SOMO→SOMO excitations leading to a computation and storage requirement of $N \times N$ matrices of dimension $(f_{NS} \times f_{NS})$. Additional cost reduction is achieved by making use of the fact that every coupling coefficient matrix that corresponds to a DOMO→Virtual excitation can be obtained by transposing the matrix belonging to a SOMO→SOMO excitation. For example, the matrix associated with

$$\langle 22220; 0|E_5^4|22211; 0\rangle \tag{8}$$

corresponds to a SOMO→SOMO excitation when read from right to left while it describes a DOMO→Virtual excitation when interpreted in the opposite direction. Analogously, the coefficient matrix associated with a SOMO→Virtual excitation is connected to a DOMO→SOMO coupling coefficient matrix through transposition and a change of sign. Hence, one only needs to compute and store $(N+1) \times N$ DOMO→SOMO and $N \times N$ SOMO→SOMO coupling coefficient matrices for every pair of N and S that occurs in the set of CSFs in equation (1).

C. Replacement Operators and Permutations

The starting point of the presented recursive evaluation scheme for coupling coefficients is the realization that the action of a replacement operator onto a CSF is analogous to the action of a permutation operator onto a spin eigenfunction combined with a parity factor. This analogy can be illustrated by considering a system with $N = 4$ unpaired electrons and a total spin of $S = 0$. For this case there are $f_{NS} = 2$ degenerate spin eigenfunctions X_1 and X_2 that correspond to the branching diagrams depicted in Figure 1. According to the genealogical spin coupling scheme[27] these two functions are given by the following linear combinations of primitive spin functions

$$\begin{aligned} |X_1(N=4,S=0,M_s=0)\rangle =& \frac{1}{\sqrt{3}}|\alpha\alpha\beta\beta\rangle - \frac{1}{\sqrt{12}}|\alpha\beta\alpha\beta\rangle - \frac{1}{\sqrt{12}}|\alpha\beta\beta\alpha\rangle \\ & - \frac{1}{\sqrt{12}}|\beta\alpha\alpha\beta\rangle - \frac{1}{\sqrt{12}}|\beta\alpha\beta\alpha\rangle + \frac{1}{\sqrt{3}}|\beta\beta\alpha\alpha\rangle \end{aligned} \tag{9a}$$

$$|X_2(4,0,0)\rangle = \frac{1}{\sqrt{4}}|\alpha\beta\alpha\beta\rangle - \frac{1}{\sqrt{4}}|\alpha\beta\beta\alpha\rangle - \frac{1}{\sqrt{4}}|\beta\alpha\alpha\beta\rangle + \frac{1}{\sqrt{4}}|\beta\alpha\beta\alpha\rangle \tag{9b}$$

Throughout the current work only spin eigenfunctions with $M_s = S$ will be targeted and

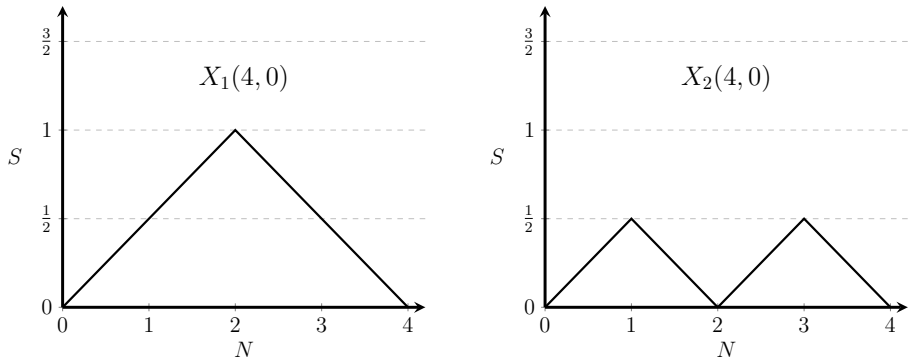

FIG. 1. Branching diagrams corresponding to two genealogical spin eigenfunctions for $N = 4$ and $S = 0$.

thus the M_s label is dropped in the following. For the current argument we will further assume that the four unpaired electrons are located in four molecular orbitals ϕ_2, ϕ_3, ϕ_4 and ϕ_5 while an additional molecular orbital, ϕ_1, is doubly occupied. Using X_1 and X_2 two orthonormal spin-adapted configuration state functions can be effortlessly constructed for

this orbital configuration:

$$|\Phi_1^{21111}(4,0)\rangle = \frac{1}{\sqrt{3}}|1\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|1\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|1\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|1\bar{1}2\bar{3}4\bar{5}\rangle \tag{10a}$$
$$- \frac{1}{\sqrt{12}}|1\bar{1}2\bar{3}4\bar{5}\rangle + \frac{1}{\sqrt{3}}|1\bar{1}2\bar{3}4\bar{5}\rangle$$
$$|\Phi_2^{21111}(4,0)\rangle = \frac{1}{\sqrt{4}}|1\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{4}}|1\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{4}}|1\bar{1}2\bar{3}4\bar{5}\rangle + \frac{1}{\sqrt{4}}|1\bar{1}2\bar{3}4\bar{5}\rangle \tag{10b}$$

A bar over the orbital label in equations (10a) and (10b) indicates occupation with a β electron and the exponent of Φ denotes the orbital occupation pattern $\mathbf{n}$. Note, that every ket represents a vector in Fock space, i.e. an antisymmetrized orbital product. The action of E_1^4 on Φ_1 is now given by

$$E_1^4|\Phi_1^{21111}(4,0)\rangle = \frac{1}{\sqrt{3}}|4\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|1\bar{4}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|4\bar{1}2\bar{3}4\bar{5}\rangle - \frac{1}{\sqrt{12}}|1\bar{4}2\bar{3}4\bar{5}\rangle \tag{11a}$$
$$- \frac{1}{\sqrt{12}}|4\bar{1}2\bar{3}4\bar{5}\rangle + \frac{1}{\sqrt{3}}|1\bar{4}2\bar{3}4\bar{5}\rangle$$
$$= -\frac{1}{\sqrt{3}}|\bar{1}234\bar{4}\bar{5}\rangle + \frac{1}{\sqrt{12}}|12\bar{3}\bar{4}\bar{4}\bar{5}\rangle + \frac{1}{\sqrt{12}}|\bar{1}2\bar{3}\bar{4}\bar{4}5\rangle + \frac{1}{\sqrt{12}}|1\bar{2}3\bar{4}\bar{4}\bar{5}\rangle \tag{11b}$$
$$+ \frac{1}{\sqrt{12}}|\bar{1}\bar{2}34\bar{4}\bar{5}\rangle - \frac{1}{\sqrt{3}}|1\bar{2}\bar{3}\bar{4}\bar{4}5\rangle$$
$$= C_1|\Phi_1^{11121}\rangle + C_2|\Phi_2^{11121}\rangle \tag{11c}$$

A comparison of the occupation pattern in the singly occupied parts of equation (11b) with equation (10a) reveals that the action of E_1^4 on Φ_1 is equivalent to the action of $\hat{P}^{N=4}_{(3124)}$ on $X_1(4,0)$ with a parity factor of -1 since

$$-\hat{P}^{N=4}_{(3124)}|X_1(N,S)\rangle = -\frac{1}{\sqrt{3}}|\beta\alpha\alpha\beta\rangle + \frac{1}{\sqrt{12}}|\alpha\alpha\beta\beta\rangle + \frac{1}{\sqrt{12}}|\beta\alpha\beta\alpha\rangle + \frac{1}{\sqrt{12}}|\alpha\beta\alpha\beta\rangle \tag{12}$$
$$+ \frac{1}{\sqrt{12}}|\beta\beta\alpha\alpha\rangle - \frac{1}{\sqrt{3}}|\alpha\beta\beta\alpha\rangle$$

Thus the coefficients in equation (11c) can be determined through finding the representation of permutation operators in the basis of spin eigenfunctions, i.e.

$$C_1 = -\langle X_1(4,0)|\hat{P}^{N=4}_{(3124)}|X_1(4,0)\rangle \tag{13a}$$
$$C_2 = -\langle X_2(4,0)|\hat{P}^{N=4}_{(3124)}|X_1(4,0)\rangle \tag{13b}$$

For a general DOMO→SOMO excitation that corresponds to a replacement operator E_q^p the relative position of orbitals p and q, depicted in Figure 2 and henceforth denoted as p^{rel} and q^{rel}, determine the corresponding permutation $\hat{P}^N$ through

$$\hat{P}^N(E_q^p) = \begin{cases} \hat{1}^N & \text{if } p^{rel} = q^{rel} \\ \hat{1}^N & \text{if } p^{rel}+1 = q^{rel} \\ \hat{P}^N_{(1\cdots(q^{rel}-1)p^{rel}q^{rel}(q^{rel}+1)\cdots(p^{rel}-1)(p^{rel}+1)\cdots N)} & \text{if } p^{rel} > q^{rel} \\ \hat{P}^N_{(1\cdots(p^{rel}-1)(p^{rel}+1)\cdots(q^{rel}-1)p^{rel}q^{rel}(q^{rel}+1)\cdots N)} & \text{if } q^{rel} > p^{rel}+1 \end{cases} \tag{14}$$

Furthermore, the difference between the relative positions determines the parity factor as it contains the essential information about the number of pairwise exchanges of singly occupied orbitals that is required to establish the standard orbital order after the application of E_q^p on a given CSF (cf. the step from equation (11a) to equation (11b)).

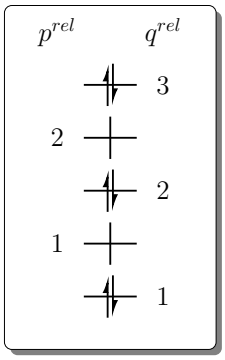

FIG. 2. All unique relative positions of the donor and acceptor orbitals relative to the set of singly occupied orbitals for the case of $N = 2$ unpaired electrons.

In case of a SOMO→SOMO excitation the connection to a permutation operator is not quite as obvious as for DOMO→SOMO excitations. At the heart of the problem lies the fact that a SOMO→SOMO excitation connects CSFs with different numbers of unpaired electrons: When an operator E^p_q acts on a CSF Φ_I with N unpaired electrons and total spin S and both, p and q, are singly occupied then the resulting set of CSFs Φ_J feature only $N-2$ unpaired electrons with the same total spin. The simplest situation of that kind is met when the orbitals in a given $\Phi_I(S,N)$ are arranged such that the electron spins of the singly occupied orbitals p and q take positions $N-1$ and N in the corresponding spin eigenfunction $X_k(N,S)$. For example, the pair of $p = 5$ and $q = 4$ fulfills this requirement for $|\Phi_1^{21111}(4,0)\rangle$ and $|\Phi_2^{21111}(4,0)\rangle$. In this case we find

$$E^5_4\,|\Phi_1^{21111}(4,0)\rangle = 0\,|\Phi_1^{21102}(2,0)\rangle \tag{15a}$$

$$E^5_4\,|\Phi_2^{21111}(4,0)\rangle = \sqrt{2}\,|\Phi_1^{21102}(2,0)\rangle \tag{15b}$$

Note, that there exists only a single spin eigenfunction $X(2,0)$ and accordingly only a single CSF with occupation pattern 21102 and $S = 0$. Since the spin eigenfunctions used to construct the CSFs here follow the genealogical coupling scheme the coefficients on the right hand side of equations (15a) and (15b) can be expressed as sum over Clebsch-Gordan coefficients. In the following we will establish this connection for the general case of N unpaired electrons.

Within the genealogical coupling scheme the set of spin eigenfunctions $\{X_k(N,S,M_s)\}$ are constructed from spin eigenfunctions with $N-2$ through coupling with two additional electron spins. Figure 3 depicts all possible routes through a branching diagram that connect the f_{NS} spin eigenfunctions with N and S to spin eigenfunctions with $N-2$ unpaired electrons and total spins of $S-1$, S and $S+1$. The f_{11} first eigenfunctions correspond to antiferromagnetic coupling of the two spins associated with the $N-1$'th and N'th electron to the f_{11} spin eigenfunctions $\{X_{k'}(N-2,S+1)\}$. Then, f_{12} and f_{21} spin eigenfunctions originate from *up-down* and *down-up* paths while the last f_{22} spin eigenfunctions are constructed by ferromagnetically coupling the $N-1$'th and N'th electron spin to f_{22} spin eigenfunctions $\{X_{k'}(N-2,S-1)\}$. The Clebsch-Gordan coefficients that correspond to the four pathways are given by[27]

$$\begin{aligned} X_k(N,S,S) =& \frac{\sqrt{2}}{\sqrt{(2S+2)(2S+3)}} X_{k'}(N-2,S+1,S-1)\times[\alpha(N-1)\alpha(N)] \\ &- \frac{\sqrt{2S+1}}{\sqrt{(2S+2)(2S+3)}} X_{k'}(N-2,S+1,S)\times[\alpha\beta+\beta\alpha] \\ &+ \frac{\sqrt{(2S+1)(2S+2)}}{\sqrt{(2S+2)(2S+3)}} X_{k'}(N-2,S+1,S+1)\times[\beta\beta] \end{aligned} \tag{16}$$

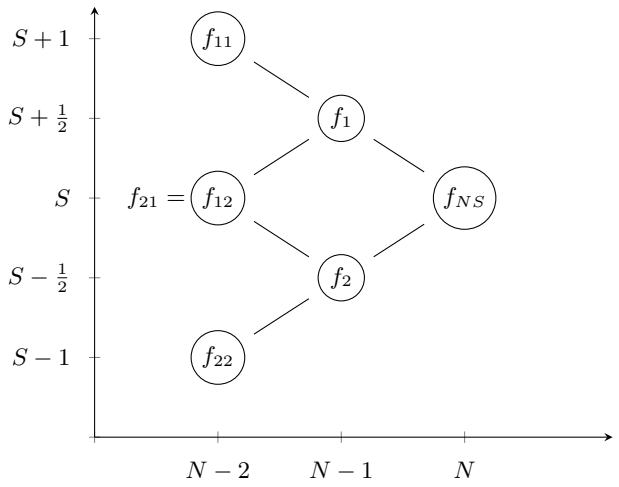

FIG. 3. Possible pathways towards the f_{NS} spin eigenfunctions with N unpaired electrons and a total spin S from spin eigenfunctions with $N-2$ unpaired electrons and total spins of $S-1$, S and $S+1$. For each point in the branching diagram the spin degeneracy is given.

where $k = k' = 1 \cdots f_{11}$ and

$$\begin{aligned} X_k(N,S,S) = &- \frac{\sqrt{2S}}{\sqrt{(2S+1)(2S+2)}} X_{k'}(N-2,S,S-1) \times [\alpha\alpha] \\ &- \frac{1}{\sqrt{(2S+1)(2S+2)}} X_{k'}(N-2,S,S) \times [\beta\alpha] \\ &+ \frac{2S+1}{\sqrt{(2S+1)(2S+2)}} X_{k'}(N-2,S,S) \times [\alpha\beta] \end{aligned} \tag{17}$$

where $k' = 1 \cdots f_{12}$ and $k = f_{11} + k'$. Furthermore,

$$\begin{aligned} X_k(N,S,S) = &- \frac{\sqrt{2S}}{\sqrt{2S(2S+1)}} X_{k'}(N-2,S,S-1) \times [\alpha\alpha] \\ &+ \frac{2S}{\sqrt{2S(2S+1)}} X_{k'}(N-2,S,S) \times [\beta\alpha] \end{aligned} \tag{18}$$

where $k' = 1 \cdots f_{21}$ and $k = f_{11} + f_{12} + k'$ and

$$X_k(N,S,S) = X_{k'}(N-2,S-1,S-1) \times [\alpha\alpha] \tag{19}$$

where $k' = 1 \cdots f_{22}$ and $k = f_{11} + f_{12} + k'$. When p and q take positions N and $N-1$, the action of E_q^p on a CSF $\Phi_k^{\mathbf{n}}(N,S)$ is to convert $[\alpha(N-1)\beta(N)]$ as well as $-[\beta(N-1)\alpha(N)]$ to $[\alpha(N)\beta(N)]$ while all other combinations vanish. Therefore, the coefficients $\{C_{k'}^k\}$ in

$$E_q^p \, |\Phi_k^{\mathbf{n}}(N,S)\rangle = \sum_{k'} C_{k'}^k \, |\Phi_{k'}^{\mathbf{n}'}(N-2,S)\rangle \tag{20}$$

simply become

$$C_{k'}^k = \begin{cases} \sqrt{\frac{2S+2}{(2S+1)(2S+2)}} & \text{if } k = k' + f_{11} \\ \sqrt{\frac{2S}{2S(2S+1)}} & \text{if } k = k' + f_{11} + f_{12} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

As a consequence, every spin eigenfunction $X_k(N,S,S)$ that follows the *up-down* or *down-up* path in Figure 3 is connected to at most two spin eigenfunctions $X_{k'}(N-2,S,S)$. Obviously, all spin eigenfunctions $X_k(N,S,S)$ that are associated with a *down-down* or *up-up* route are not connected to any eigenfunctions $X_{k'}(N-2,S,S)$ through the action of E^p_q. Hence, the coupling coefficient matrix has a dimension of $(f_{N-2\ S} \times f_{NS})$ (see also discussion above).

If p and q do not take positions N and $N-1$, respectively, the action of the corresponding operator E^p_q on a given CSF has to be described by a two-step procedure. First, the electrons are permuted such that the larger index of p^{rel} and q^{rel} takes position N while the smaller index occupies position $N-1$. Then equation (21) is applied to determine nonvanishing coupling coefficients. For example, to describe $E^3_2\,|\Phi^{21111}_1(4,0)\rangle$ one would have to apply the combination of permutation operators

$$\hat{R}^{N=4}_{(2314)} \times \hat{P}^{N=4}_{(1342)} \tag{22}$$

bring $q^{rel}=1$ to position 3 — bring $p^{rel}=2$ to position 4

on $X_1(4,0)$ before using equation (21) to determine nonvanishing coupling coefficients. Analogous to the above discussed case of DOMO→SOMO excitations the representation matrices of the applied permutation operators have to be combined with a parity factor that depends on p^{rel} and q^{rel}.

In summary, we have established that the representation matrices of nonredundant excitation operators in the basis of CSFs can be expressed through

$$\mathbf{A}^N_S(E^p_q) = s(p,q)\cdot \mathbf{U}_S(\hat{P}^N) \qquad \text{(DOMO} \rightarrow \text{SOMO)} \tag{23}$$

$$\mathbf{A}^N_S(E^p_q) = s(p,q)\cdot \mathbf{C} \times \mathbf{U}_S(\hat{R}^N) \times \mathbf{U}_S(\hat{P}^N) \qquad \text{(SOMO} \rightarrow \text{SOMO)} \tag{24}$$

where $s(p,q)=1,-1$ is a sign factor, $\mathbf{C}$ is the matrix of projection coefficients from equation (21) while $\mathbf{U}_S(\hat{P}^N)$ and $\mathbf{U}_S(\hat{R}^N)$ are representation matrices of the required permutation operators in the basis of genealogical spin eigenfunctions with N unpaired electrons and totals spin S. An efficient algorithm for the efficient evaluation of these matrices is described in the following section.

D. Recursive Formulation of Representation Matrices

The presented algorithm relies on a recursive scheme for the construction of representation matrices $\mathbf{U}_S(\hat{P}^N)$ from $(N-1)$ - electron permutation operators that was originally developed by Yamanouchi.[28] Kotani later elaborated on this method and presented it in a clear way.[29] This work is based on the description of the method given by Pauncz.[27] In the framework of Yamanouchi's method, any permutation is expressed as product of permutations that do not affect the N'th electron and the operator $(N-1,\ N)$ that exchanges electrons $N-1$ and N.

The representation matrix of any permutation $\hat{P}^N$ of N electrons that does not affect the N'th electron in the basis of spin eigenfunctions with N and S can be readily constructed when the representation matrix of the same permutation for $N-1$ and $S+\frac{1}{2}$ as well as $S-\frac{1}{2}$ is already known; it reads

$$\mathbf{U}_S(\hat{P}^N) = \begin{pmatrix} \mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1}) & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{S-\frac{1}{2}}(\hat{P}^{N-1}) \end{pmatrix} \tag{25}$$

where $\mathbf{U}_S(\hat{P}^N)$ is a $(f_{NS} \times f_{NS})$ matrix while $\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$ and $\mathbf{U}_{S-\frac{1}{2}}(\hat{P}^{N-1})$ are square matrices of dimension $f_1 = f_{11} + f_{12}$ and $f_2 = f_{21} + f_{22}$, respectively. Furthermore, the representation matrix of $(N-1,\ N)$ can be readily formulated when the two spin labels for electrons $N-1$ and N are interchanged in equations (16) through (19) and the resulting functions are projected on the original set of spin eigenfunctions:

$$\mathbf{U}_S(N-1,\ N) = \begin{pmatrix} \mathbf{1}_{f_{11}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -a\mathbf{1}_{f_{12}} & b\mathbf{1} & \mathbf{0} \\ \mathbf{0} & b\mathbf{1} & a\mathbf{1}_{f_{21}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{f_{22}} \end{pmatrix} \tag{26}$$

Here, $\mathbf{1}$ denotes unity matrices of dimension f_{11}, f_{12}, f_{21} and f_{22}, respectively, while

$$a = \frac{1}{2S+1} \tag{27}$$

$$b = \sqrt{1-a^2} \tag{28}$$

are simple constants.

In the case of DOMO→SOMO excitations for N and S the representation matrices of the corresponding permutations can be expressed through

a) equation 25 if $\hat{P}^N(N) = N$ which means that electron N is not affected by $\hat{P}^N$

b) $\mathbf{U}_S(\hat{P}^N) = \mathbf{U}_S(\hat{Q}^N) \times \mathbf{U}_S(N-1,\ N)$, if $\hat{P}^N(N) = N-1$ which means that $\hat{P}^N$ acts to put electron N to position $N-1$.

c) $\mathbf{U}_S(\hat{P}^N) = \mathbf{U}_S(N-1,\ N) \times \mathbf{U}_S(\hat{Q}^N)$, if $\hat{P}^N(N) < N-1$ which means that $\hat{P}^N$ acts to put electron N to a position smaller than $N-1$.

The permutation $\hat{Q}^N$ that appears in cases b) and c) does not affect electron N but has a similar cyclic character as $\hat{P}^N$. Therefore, $\mathbf{U}_S(\hat{Q}^N)$ can be constructed according to equation (25) and the descending permutation $\hat{Q}^{N-1}$ can be treated according to cases a) through c) again where N has been reduced by one. Accordingly, the representation matrix $\mathbf{U}_S(\hat{P}^N)$ of any permutation that corresponds to a DOMO→SOMO excitation may be computed by recursively applying cases a) through c) until $N = 2$ is reached in which case the representation matrices take a simple form.

The representation matrices for the two permutations that are required to express the action of a given SOMO→SOMO excitation can be computed in a similar fashion. However, subtle differences occur which necessitate a different case structure for their evaluation. In general, every involved permutation either promotes a given electron with index p^{rel} to position N or $N-1$ (see above). The relative positioning of all other electrons remains unchanged. Accordingly, the following case structure can be applied to compute the corresponding representation matrices:

a) Apply equation 25 if p^{rel} is promoted to $N-1$

b) $\mathbf{U}_S(\hat{P}^N) = \mathbf{U}_S(N-1,\ N) \times \mathbf{U}_S(\hat{Q}^N)$, if p^{rel} is promoted to N

c) $\mathbf{U}_S(\hat{P}^N) = \mathbf{1}$, if $p^{rel} = N$

Here, permutation $\hat{Q}^N$ promotes electron p^{rel} to position $N-1$. Therefore its corresponding representation matrix can be constructed according to equation (25) while its descendant permutations have to be treated again according to case b) with N being reduced by one. Again, these relations allow for a simple recursive construction of all required representation matrices.

III. IMPLEMENTATION AND PERFORMANCE

A. Generation of Coupling Coefficient Matrices

Algorithm 1: Algorithm for the recursive construction of DOMO→SOMO coupling coefficient matrices.

Function *GenerateRepresentationMatrix(N,S,*$\hat{P}^N$*,*$\mathbf{U}_S(\hat{P}^N)$*)*:
- **if** N *or* S *is invalid* **then**
  - return;
- **else if** $N = 2$ **then**
  - Construct $\mathbf{U}_S(\hat{P}^N)$ explicitly;
- **else if** $\hat{P}^N(N) \neq N$ **then**
  - Determine $\hat{Q}^{N-1}$;
  - GenerateRepresentationMatrix($N-1$,$S+\frac{1}{2}$,$\hat{Q}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{Q}^{N-1})$);
  - GenerateRepresentationMatrix($N-1$,$S-\frac{1}{2}$,$\hat{Q}^{N-1}$,$\mathbf{U}_{S-\frac{1}{2}}(\hat{Q}^{N-1})$);
  - Construct $\mathbf{U}_S(\hat{Q}^N)$ from $\mathbf{U}_{S+\frac{1}{2}}(\hat{Q}^{N-1})$ and $\mathbf{U}_{S-\frac{1}{2}}(\hat{Q}^{N-1})$;
  - Construct $\mathbf{U}_S(N-1,\ N)$;
  - **if** $\hat{P}^N(N) = N-1$ **then**
    - $\mathbf{U}_S(\hat{P}^N) = \mathbf{U}_S(N-1,\ N) \times \mathbf{U}_S(\hat{Q}^N)$;
  - **else**
    - $\mathbf{U}_S(\hat{P}^N) = \mathbf{U}_S(\hat{Q}^N) \times \mathbf{U}_S(N-1,\ N)$;
- **else**
  - Determine $\hat{P}^{N-1}$;
  - GenerateRepresentationMatrix($N-1$,$S+\frac{1}{2}$,$\hat{P}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$);
  - GenerateRepresentationMatrix($N-1$,$S-\frac{1}{2}$,$\hat{P}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$);
  - Construct $\mathbf{U}_S(\hat{P}^N)$ from $\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$ and $\mathbf{U}_{S-\frac{1}{2}}(\hat{P}^{N-1})$;

Function *Main*:
- Determine $\hat{P}^N$ that corresponds to E^p_q;
- GenerateRepresentationMatrix(N,S,$\hat{P}^N$,$\mathbf{U}_S(\hat{P}^N)$);
- Determine sign s that corresponds to p^{rel} and q^{rel};
- Compute coupling coefficient matrix $\mathbf{A}^N_S(E^p_q) = s \cdot \mathbf{U}_S(\hat{P}^N)$

The recursive scheme for the construction of permutation representation matrices in the basis of spin eigenfunctions and in turn the generation of one-electron coupling coefficent matrices is amenable for an efficient implementation in computer code. Algorithms 1 and 2 provide pseudocode for the implementations in the `recursiveCC` standalone program as well as the full-CI module of the MOLBLOCK program (see below).[30,31] The former can be obtained (also as shared library) free of charge at our website[1] while the latter will be made publicly available in due course. Importantly, algorithms 1 and 2 only evaluate nonzero elements of every $\mathbf{U}_S(\hat{P}^N)$ explicitly and hence no computational effort is wasted on calculating zeros. Furthermore, the implementations require a minimal amount of memory as the descending matrices $\mathbf{U}_{S\pm\frac{1}{2}}(\hat{P}^{N-1})$ and $\mathbf{U}_{S\pm\frac{1}{2}}(\hat{Q}^{N-1})$ are not generated as separate entities but as part of the original matrix $\mathbf{U}_S(\hat{P}^N)$. This strategy is feasible because every entry of $\mathbf{U}_S(\hat{P}^N)$ remains constant after it has been evaluated once. Using a desktop PC with an Intel® Core™ i9-10900 with 2.80 GHz and 32GB RAM the computation of all nonredundant coupling coefficient matrices for an active space with 20 electrons

[1] `www.chemie.hu-berlin.de/en/forschung-en/theoretical-chemistry/downloads`

in 20 orbitals and a total spin of $S = 0$ required only 6 h and 46 min. During these test calculations the computer time was dominated by the final matrix multiplication step of the SOMO→SOMO algorithm which took 92% of the computer time. In contrast to the efficient recursive generation presented here, a creation of coupling coefficient matrices utilizing a straightforward contraction of transformation matrices is only feasible for up to 16 unpaired electrons.[9] In our test calculations using the ORCA program package[32] in its version 5.0.1 the generation of all coupling coefficient prototypes for a (16,16) active space required ca. 14 h. In comparison, the creation of one-electron coupling coefficent matrices with `recursiveCC` took only 13 s which corresponds to a speedup of about four orders of magnitude.

Algorithm 2: Algorithm for the recursive construction of SOMO→SOMO coupling coefficient matrices.

Function *GenerateRepresentationMatrix(N,S,p^{rel},$\hat{P}^N$,$\mathbf{U}_S(\hat{P}^N)$)*:
- **if** *N or S is invalid* **then**
  - return;
- **else if** $N = 2$ **then**
  - Construct $\mathbf{U}_S(\hat{P}^N)$ explicitly;
- **else if** $p = N$ **then**
  - $\mathbf{U}_S(\hat{P}) = \mathbf{1}$
- **else if** $\hat{P}^N(N) \neq N$ **then**
  - Determine $\hat{Q}^{N-1}$;
  - GenerateRepresentationMatrix($N-1$,$S+\frac{1}{2}$,p^{rel},$\hat{Q}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{Q}^{N-1})$);
  - GenerateRepresentationMatrix($N-1$,$S-\frac{1}{2}$,p^{rel},$\hat{Q}^{N-1}$,$\mathbf{U}_{S-\frac{1}{2}}(\hat{Q}^{N-1})$);
  - Construct $\mathbf{U}_S(\hat{Q}^N)$ from $\mathbf{U}_{S+\frac{1}{2}}(\hat{Q}^{N-1})$ and $\mathbf{U}_{S-\frac{1}{2}}(\hat{Q}^{N-1})$;
- **else**
  - Determine $\hat{P}^{N-1}$;
  - GenerateRepresentationMatrix($N-1$,$S+\frac{1}{2}$,p^{rel},$\hat{P}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$);
  - GenerateRepresentationMatrix($N-1$,$S-\frac{1}{2}$,p^{rel},$\hat{P}^{N-1}$,$\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$);
  - Construct $\mathbf{U}_S(\hat{P}^N)$ from $\mathbf{U}_{S+\frac{1}{2}}(\hat{P}^{N-1})$ and $\mathbf{U}_{S-\frac{1}{2}}(\hat{P}^{N-1})$;

Function *Main*:
- Determine $\hat{P}^N$ that corresponds to moving $max(p^{rel}, q^{rel})$ to position N;
- GenerateRepresentationMatrix(N,S,$max(p^{rel}, q^{rel})$,$\hat{P}^N$,$\mathbf{U}_S(\hat{P}^N)$);
- Determine $\hat{R}^N$ that corresponds to moving $min(p^{rel}, q^{rel})$ to position $N-1$;
- GenerateRepresentationMatrix(N,S,$min(p^{rel}, q^{rel})$,$\hat{R}^N$,$\mathbf{U}_S(\hat{R}^N)$);
- Determine sign s that corresponds to p^{rel} and q^{rel};
- Construct projection matrix $\mathbf{C}$ according to equation (21);
- Compute coupling coefficient matrix $\mathbf{A}_S^N(E_q^p) = s \cdot \mathbf{C} \times \mathbf{U}_S(\hat{R}^N) \times \mathbf{U}_S(\hat{P}^N)$

Since the computing time for the different coupling coefficient matrices is almost perfectly evenly distributed the parallel usage of the presented algorithm scales well with the number of processors. As demonstrated in Figure 4, a simple message passing interface (MPI) based parallel implementation lead to almost perfectly linear scaling of the speedup for the calculation of coupling coefficients for up to 18 unpaired electrons on up to 10 processors which was the maximum tested. Nevertheless, we would like to mention that the factorial scaling of the required computational effort and the corresponding storage requirement will quickly impede calculations with many more than 20 unpaired electrons and $S = 0$. Therefore, systems whose description necessitate CSFs with significantly more unpaired electrons are best tackled by stochastic or selected CI variants that rely on the selection of single CSFs rather than entire configurations.[9,19]

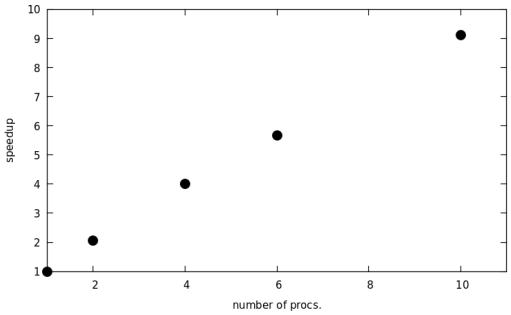

FIG. 4. Speedup for the calculation of coupling coefficient matrices for up to 18 unpaired electrons with respect to the number of used processors with a simple MPI based parallel implementation of the presented algorithm.

B. Spin-Adapted Full-CI

When dealing with CI expansions that involve CSFs with many unpaired electrons, storage of the coupling coefficient matrices will become a memory bottleneck on account of the steep scaling of the space requirement. For example, when 20 unpaired electrons occur, the number of nonredundant SOMO→SOMO coupling coefficient matrices amounts to 400 with each matrix requiring around 2 GB of space. An efficient approach to reduce the required memory by decomposing any coupling coefficient matrix into a product of two matrices which are associated with a single orbital label has been outlined by Knowles and Werner.[21] Yet, this procedure doubles the amount matrix-times-matrix operations during the σ-vector generation and still requires a significant amount of memory when large coupling coefficient matrices are being contracted. An alternative strategy for the σ-vector generation that aims to minimize the storage requirements is to construct and use large coupling coefficient matrices 'on-the-fly'. In that case an efficient means to calculate entire coupling coefficient matrices as a whole –like the one presented here– becomes absolutely vital. We have adopted this 'direct' strategy during the design of a configuration-driven and thus spin-adapted Full-CI implementation in our MOLBLOCK program.[30,31] Within our implementation the critical $\Delta\sigma$ contribution to the σ-vector is evaluated according to Siegbahn's suggestion from equation (4). Algorithm 3 outlines the basic steps to constructing the $\mathbf{D}$ matrix. The final contraction of coupling coefficient matrices with $\mathbf{ID}$ is done analogously.

Since neither the computation nor the storage of all coupling coefficient matrices for up to $N^{max} = 16$ unpaired electrons is particularly demanding they are calculated once and kept in memory. Then all one-electron replacement connections between configurations $E^p_q \left|\mathbf{n}_I\right\rangle \rightarrow \left|\mathbf{n}_J\right\rangle$ with $N \leq N^{max}$ are stored in a simple list L_1 alongside their essential information, i.e N, T, p^{rel}, q^{rel} where T is the excitation type (SOMO→SOMO etc.). Finally, the contributions of the configurations with $N \leq N^{max}$ to $\mathbf{D}$ are computed by looping over the connections and contracting the required coupling coefficient matrix with $\mathbf{C}_J$, the $f_{N_J S}$-dimensional part of the trial vector $\mathbf{C}$ corresponding to $\mathbf{n}_J$. For configurations with $N > N^{max}$ the algorithm is adapted as indicated above. Importantly, the list L_2 that holds the connections between these configurations is ordered according to a key with elements N, T, p^{rel}, q^{rel} that uniquely defines the involved coupling coefficient matrix. Thus, for every unique coupling coefficient matrix $\mathbf{A}^N_S(E^{p^{rel}}_{q^{rel}})$ the list L_2 holds another list L_3 of connections that $\mathbf{A}^N_S(E^{p^{rel}}_{q^{rel}})$ is involved in. This list structure allows us to loop over unique coupling coefficient matrices first and construct them 'on-the-fly' before looping over the associated connections and finally evaluating the contribution to $\mathbf{D}$ through contraction of $\mathbf{A}^N_S(E^{r^{el}}_{q^{rel}})$ and $\mathbf{C}_J$.

Algorithm 3: Algorithm for the computation of the **D**-matrix from equation (4)

Function *Compute* **D:**
- Create list of configurations $\{\mathbf{n}_I\}$;
- Calculate and store all $\mathbf{A}_S^N(E_{q^{rel}}^{p^{rel}})$ for $N \leq 16$;
- Create list of connections $L_1 = \{(N, T, p^{rel}, q^{rel}), p, q, \mathbf{n}_I, \mathbf{n}_J\}$ for $N \leq N^{max}$;
- Create list of connections $L_2 = \{(N, T, p^{rel}, q^{rel}) \rightarrow L_3 = \{\mathbf{n}_I, \mathbf{n}_J, p, q\}\}$ for $N > N^{max}$;
- **for** *elements in* L_1 **do**
  - Get coupling coefficient matrix $\mathbf{A}_S^N(E_{q^{rel}}^{p^{rel}})$;
  - Form $\mathbf{D}_{pq} = \mathbf{A}_S^N(E_{q^{rel}}^{p^{rel}})\mathbf{C}_J$;
- **for** *elements in* L_2 **do**
  - Evaluate coupling coefficient matrix $\mathbf{A}_S^N(E_{q^{rel}}^{p^{rel}})$;
  - **for** *elements in* $L_3(N, T, p^{rel}, q^{rel})$ **do**
    - Form $\mathbf{D}_{pq} = \mathbf{A}_S^N(E_{q^{rel}}^{p^{rel}})\mathbf{C}_J$;

To test the performance of the presented algorithms, a series of CASCI calculations was run on benzene with the def2-SVP basis set[33] and different active spaces comprising n electrons in m active orbitals, henceforth denoted as CASCI(n,m). In terms of their computing times the following steps are relevant: Recursive construction of all coupling coefficient matrices for up to N^{max}=16 unpaired electrons (t(CC)) and configurations (t(CFG)), creation of connections lists t(conn.), σ-vector evaluation ($t(\sigma)$) and the linear algebra operations related to the Davidson diagonalization routine (t(lin. alg.)). Table I summarizes the timings for every step as obtained using up to 12 cores of a single Intel® Xeon(R) Gold 6128 CPU with 3.40 GHz and 1.5 TB RAM. It should be noted that neither the recursive coupling coefficient build for up to N^{max} unpaired electrons utilized in our Full-CI implementation nor the routine to generate all configurations has been parallelized since none of the two steps constitutes a computational bottleneck. Moreover, for σ-vector evaluations, only the average time for a single vector build is listed in Table I to make the data independent of the number of steps required to achieve convergence of the underlying Davidson diagonalization routine.[24] In the present case, 10 iterations were sufficient to reach convergence. For technical reasons the first iteration invokes the σ-vector build twice. Therefore, $t(\sigma)$ for the first iteration was divided by 2 during the averaging procedure.

Obviously, the σ-vector generation is the most time-consuming step in the presented example with every σ-vector build taking 241.22 s on average in a serial run of the code. Yet, the time for this step can be considerably reduced by running the code in parallel. When 12 cores are used the time for a single σ-vector build reduces to 44.56 s. Likewise, t(conn.) and t(lin. alg.) reduce from 155.14 s to 23.69 s and from 485.66 s to 73.42 s, respectively. The total run time for the CASCI calculation reduces from 3365.64 s to 654.96 s. It is noteworthy that these timings are on the same order of magnitude as the run times reported recently by Fales *et al.* for a spin-adapted CASCI(16,16) calculation on ethylene.[11] Yet, a direct comparison of these timings can be misleading since

i) the number of CI iterations and accordingly the number of σ builds and linear algebra operations is not equal in the two calculations (11 vs 21).

ii) the computer hardware used in both cases greatly differ and thus do not allow for a straightforward comparison. More precisely, the calculation reported in reference 11 employed an Intel® Xeon(R) E5-2643@3.40 GHz CPU together with a NVIDIA V100 graphical processing unit (GPU).

TABLE I. Relevant timings (in s) for a CASCI(16,16) calculation on benzene.

#(cores)	t**(CC)**	t**(CFG)**	t**(conn.)**	t**(lin. alg.)**[a]	$t(\sigma)$[b]
			Required 1 ×		Required 10 ×
1	55.63	15.68	155.14	485.66	241.22
2	50.29	14.16	102.36	275.60	166.81
4	51.10	11.88	56.72	153.62	91.67
8	55.28	11.09	33.43	93.17	56.30
12	56.97	10.71	23.69	73.42	44.56

[a] total sum of time of linear algebra operations related to Davidson routine
[b] time for single σ-vector builds were averaged over all iterations.

To demonstrate the effect that calculating coupling coefficient matrices 'on-the-fly' has on the computing times the same calculation with a reduced N^{max} of 14 was run. As a result, t(CC) was reduced to 31.71 s while $t(\sigma)$ increased to 344.54 s. All other timings remained similar as compared to when N^{max} was set to 16. With 12 cores $t(\sigma)$ reduced to 99.15 s.

Of course, further enlargement of the active space size leads to a drastic increase of the required computer time. For example, a CASCI calculation of the benzene cation radical with a (17,17) active space ($S = \frac{1}{2}$) required 71601.75 s using a single core. This time distributes as follows over the aforementioned steps: t(CC) = 1154.35 s, t(CFG) = 62.57 s, t(conn.) = 658.36 s, t(lin. alg.) = 6321.22 and $t(\sigma) \times 13$ = 63405.25 s. With increasing active space size both, the time to build σ-vectors as well as their storage requirement, will eventually become prohibitive. On computers available to us, this point is met before the recursive construction of coupling coefficient matrices becomes unfeasible (see above).

IV. CONCLUSIONS

We have introduced a recursive computational approach to the efficient evaluation of one-electron coupling coefficients as they are required during electronic structure calculations of the configuration interaction type. The approach relies on the equivalence of the representation matrix of excitation operators in the basis of configuration state functions and the representation matrix of permutation operators in the basis of spin eigenfunctions. While for DOMO→SOMO excitations this connection is straightforward, SOMO→SOMO excitations in general require a combination of two permutation operators and a projection operator. With the permutation and projection operator identities at hand, the former are generated by a recursive scheme originally introduced by Yamanouchi and Kotani. On the basis of this scheme we have implemented an efficient computer code that allows the evaluation of all nonredundant coupling coefficients for systems with 20 unpaired electrons and a total spin of $S = 0$ within only a few hours on a simple Desktop-PC. Furthermore, a full-CI implementation that utilizes the presented method is shown to perform well in terms of computational timings for CASCI calculations with comparably large active spaces. More importantly, however, this work paves the way to spin-adapted and configuration driven selected configuration interaction calculations on systems with many unpaired electrons. Finally, we would like to emphasize that the code is available free of charge upon request and can thus be used by anyone to easily spin-adapt existing CI codes or write such a code from scratch.

V. DATA AVAILABILITY STATEMENT

The presented code and data described in this study are available from the corresponding author upon reasonable request.

VI. FUNDING INFORMATION

This work was funded by the Deutsche Forschungsgemeinschaft through grant RO 5688/1-1.

ACKNOWLEDGMENTS

MU and MR thank the Deutsche Forschungsgemeinschaft for funding through Emmy-Noether grant RO 5688/1-1.

REFERENCES

[1]H. Lischka, D. Nachtigallovái, A. J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti, "Multireference approaches for excited states of molecules," Chem. Rev. **118**, 7293–7361 (2018), https://doi.org/10.1021/acs.chemrev.8b00244.
[2]A. Khedkar and M. Roemelt, "Modern multireference methods and their application in transition metal chemistry," Phys. Chem. Chem. Phys. **23**, 17097–17112 (2021).
[3]P. J. Knowles and N. C. Handy, "A new determinant-based full configuration interaction method," Chem.. Phys. Lett. **111**, 315–321 (1984).
[4]J. Olsen, B. Roos, P. Jørgensen, and H. Jensen, "Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces," J. Chem. Phys. **89**, 2185 (1988).
[5]J. Olsen, P. Jørgensen, and J. Simons, "Passing the one-billion limit in full configuration-interaction (fci) calculations," Chem. Phys. Lett. **169**, 463–472 (1990).
[6]S. Evangelisti, G. L. Bendazzoli, R. Ansaloni, F. Durì, and E. Rossi, "A one billion determinant full ci benchmark on the cray t3d," Chem. Phys. Lett. **252**, 437–446 (1996).
[7]B. S. Fales and B. G. Levine, "Nanoscale multireference quantum chemistry: Full configuration interaction on graphical processing units," J. Chem. Theory Comput. **11**, 4708–4716 (2015).
[8]B. S. Fales and T. J. Martínez, "Efficient treatment of large active spaces through multi-gpu parallel implementation of direct configuration interaction," J. Chem. Theory Comput. **16**, 1586–1596 (2020).
[9]V. G. Chilkuri and F. Neese, "Comparison of many-particle representations for selected-ci i: A tree based approach," J. Comput. Chem. **42**, 982–1005 (2021).
[10]V. G. Chilkuri and F. Neese, "Comparison of many-particle representations for selected configuration interaction: Ii. numerical benchmark calculations," J. Chem. Theory Comput. **17**, 2868–2885 (2021), pMID: 33886300, https://doi.org/10.1021/acs.jctc.1c00081.
[11]B. S. Fales and T. J. Martínez, "Fast transformations between configuration state function and slater determinant bases for direct configuration interaction," J. Chem. Phys. **152**, 164111 (2020).
[12]V. G. Chilkuri, T. Applencourt, K. Gasperich, P.-F. Loos, and A. Scemama, "Chapter three - spin-adapted selected configuration interaction in a determinant basis," in *New Electron Correlation Methods and their Applications, and Use of Atomic Orbitals with Exponential Asymptotes*, Advances in Quantum Chemistry, Vol. 83, edited by M. Musial and P. E. Hoggan (Academic Press, 2021) pp. 65–81.
[13]J. Olsen, "A direct method to transform between expansions in the configuration state function and slater determinant bases," J. Chem. Phys. **141**, 034112 (2014).
[14]M. Moshinsky, *Group Theory and The Many-Body Problem* (Gordon & Breach Publishers, 1968).
[15]J. Paldus, "Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems," J. Chem. Phys. **61**, 5321–5330 (1974).
[16]I. Shavitt, "Graph theoretical concepts for the unitary group approach to the many-electron correlation problem," Int. J. Quant. Chem. **12**, 131–148 (1977).
[17]I. Shavitt, "Matrix element evaluation in the unitary group approach to the electron correlation problem," Int. J. Quant. Chem. **14**, 5–32 (1978).
[18]H. Lischka, R. Shepard, F. B. Brown, and I. Shavitt, "New implementation of the graphical unitary group approach for multireference direct configuration interaction calculations," Int. J. Quant. Chem. **20**, 91–100 (1981).

[19]W. Dobrautz, S. D. Smart, and A. Alavi, "Efficient formulation of full configuration interaction quantum monte carlo in a spin eigenbasis via the graphical unitary group approach," J. Chem. Phys. **151**, 094104 (2019).
[20]S. Rettrup, "A recursive formula for young's orthogonal representation," Chem. Phys. Lett. **47**, 59–60 (1977).
[21]P. J. Knowles and H.-J. Werner, "An efficient method for the evaluation of coupling coefficients in configuration interaction calculations," Chem. Phys. Lett. **145**, 514–522 (1988).
[22]P. Celani and H.-J. Werner, "Multireference perturbation theory for large restricted and selected active space reference wave functions," J. Chem. Phys. **112**, 5546–5557 (2000).
[23]C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," Journal of Research of the National Bureau of Standards **45**, 252–282 (1950).
[24]E. R. Davidson, "The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices," J. Comput. Phys. **17**, 87–94 (1975).
[25]M. Crouzeix, B. Philippe, and M. Sadkane, "The davidson method," SIAM Journal on Scientific Computing **15**, 62–76 (1994).
[26]P. E. M. Siegbahn, "A new direct ci method for large ci expansions in a small orbital space," Chem. Phys. Lett. **109**, 417–423 (1984).
[27]R. Pauncz, *Spin Eigenfunctions. Construction and use.* (Plenum Press, New York and London, 1979).
[28]T. Yamanouchi, "On the calculation of atomic energy levels. iv," Proc. Phys. Math. Soc. Jpn. **18**, 623–640 (1936).
[29]M. Kotani, A. Amemiya, E.Ishiguro, and T. Kimura, *Tables of Molecular Integrals*, 2nd ed. (Maruzen, 1963).
[30]M. Roemelt, V. Krewald, and D. A. Pantazis, "Exchange coupling interactions from the density matrix renormalization group and n-electron valence perturbation theory: Application to a biomimetic mixed-valence manganese complex," J. Chem. Theory Comput. **14**, 166–179 (2018).
[31]A. Khedkar and M. Roemelt, "Active space selection based on natural orbital occupation numbers from n-electron valence perturbation theory," J. Chem. Theory Comput. **15**, 3522–3536 (2019), https://doi.org/10.1021/acs.jctc.8b01293.
[32]F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, "The orca quantum chemistry program package," J. Chem. Phys. **152**, 224108 (2020).
[33]F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy," Phys. Chem. Chem. Phys. **7**, 3297–3305 (2005).