Compression of Spin-Adapted Multi-Configurational Wave Functions in Exchange-Coupled Polynuclear Spin Systems

Giovanni Li Manni,[*,†] Werner Dobrautz,[*,†] and Ali Alavi[*,†,‡]

†*Max-Planck-Institut für Festkörperforschung, Heisenbergstraße 1, 70569 Stuttgart, Germany*

‡*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*

E-mail: g.limanni@fkf.mpg.de; w.dobrautz@fkf.mpg.de; a.alavi@fkf.mpg.de

We present a protocol based on unitary transformations of molecular orbitals to reduce the number of non-vanishing coefficients of spin-adapted configuration interaction expansions. Methods that exploit the sparsity of the Hamiltonian matrix and compactness of its eigensolutions, such as the FCIQMC algorithm in its spin-adapted implementation, are well suited to this protocol. The wave function compression resulting from this approach is particularly attractive for anti-ferromagnetically coupled polynuclear spin systems, such as transition metal cubanes in bio-catalysis and, Mott and charge-transfer insulators in solid state physics. Active space configuration interaction calculations on the stretched N_2 and square N_4 compounds, the chromium dimer, and a $[Fe_2S_2]$ model system are presented as a proof-of-concept. For the Cr_2 case large and intermediate bond distances are discussed, showing that the approach is effective in cases where static and dynamic correlation are equally important. The $[Fe_2S_2]$ case shows the general applicability of the method.

1 Introduction

Poly-nuclear transition metal and f-element systems play central roles in bio-chemical processes and as building blocks of Mott and charge-transfer insulators. Understanding their electronic structure is of paramount importance to control their properties. At the atomic level, these compounds have complex electronic structures, with several unpaired electrons per metal center distributed among near-degenerate valence d (or f) orbitals. Orbital degeneracies are partially lifted by ligand-field effects, at the price of even more complex electronic structures characterized by charge-transfer excitations between metal centers and ligands (consider the super-exchange mechanism in solids as an example[1]) and degeneracies between metal and ligand orbitals. These systems also exhibit multiple quasi-degenerate low-lying spin-states whose relative order is easily altered by small external perturbations.[2] Locally (at each metal center) Hund's rules suggest that the unpaired electrons have parallel spins. However, kinetic-exchange interactions, including direct exchange and super-exchange mechanisms, favor electrons residing in adjacent metal centers to couple with anti-parallel spins, thus, inducing anti-ferromagnetism.[3–9] Computational investigations of these systems require advanced multi-configurational electronic structure methods, such as the complete active space self-consistent field approach, CASSCF.[10–14] However, for these methods even the determination of the spin of the ground state is computationally demanding, and predictions of reaction mechanisms and electronic properties are in practice limited to systems containing at most two transition metal atoms.[15–18] When studying systems with numerous unpaired and low-spin coupled electrons the limiting step is the exponential scaling of the Hilbert space size with respect to size of the chosen active space.

This limitation is exemplified by the $\{Mn_4CaO_5\}$ cluster of photosystem II. In its relaxed form, the S1 state, the cluster consists of two d^4-Mn(III) and two d^3-Mn(IV) ions. The minimal active space for this system is the CAS(14,20), consisting of the valence orbitals and electrons of the four metal centers. In the low-spin (singlet, $S = 0$) the configuration interaction (CI) vector contains $\sim 6 \times 10^9$ Slater determinants (SDs) and $\sim 1 \times 10^9$ configuration

state functions (CSFs), quickly reaching the present computational limits. A more adequate active space would also contain orbitals and electrons of the bridging oxygens, CAS(44,35), largely exceeding the current computational limits.

In recent years, a number of methods have been developed to circumvent the exponential scaling of CAS wave functions, that use algorithms such as DMRG[19–28] or FCIQMC[29–35] as CI eigensolvers. Within the framework of the novel Stochastic-CASSCF approach[36] active spaces containing up to 38 electrons and 40 orbitals have been reported.[37,38] In FCIQMC, a finite number of "*walkers*" is used to stochastically sample CAS (or FCI) wave functions and information is stored only for those SDs that are populated by walkers at the given instantaneous imaginary-time step. For a fixed number of walkers, the stochastic representation of the wave function is generally more accurate for sparse wave functions than for dense ones. Thus, it seems relevant for methods that benefit from wave function sparsity, such as FCIQMC, to ask whether techniques exist that can reduce the number of non-vanishing coefficients in CI wave functions.

The graphical unitary group approach (GUGA)[39,40] is a technique that constrains multiconfigurational wave functions to a chosen total spin, S. The method has been pioneered by Paldus, Shavitt, and others[39,41–47] and it has been used for decades in conventional MCSCF methods. Since 2011, the GUGA approach has also been adapted to generalized active space SCF wave functions (GASSCF)[48] and to the GASPT2 approach.[49] Recently, a spin-adapted version of the FCIQMC algorithm based on GUGA has also been developed in our laboratories.[50] When used in conventional CI procedures, GUGA represents the most compact way of storing CI expansions, as it contains a much smaller number of parameters (the CSF coefficients) than the ones in Slater determinant expansions. However, SD representations are more effective in direct-CI driven procedures, as the evaluation of the sigma vector, $\sigma = \mathbf{HC}$, only relies on the Slater-Condon rules and vectorization is possible.[51,52] The advantage of both expansions, Slater determinants for computing the σ vector, and CSFs for storing the wave function parameters, can be combined at the extra cost of efficient ways

of transforming the wave function between the two bases. Already in 1976, Grabenstetter[53] suggested one of such methods. This method is currently used in many chemistry software packages, including MOLCAS,[54] LUCIA[55] and DALTON.[56] More recently, an algorithm has been suggested by Olsen[57] that avoids the large spin-coupling transformation matrix and the operation count can be reduced for systems featuring a large number of low-spin coupled unpaired electrons.

Within the GUGA formalism one additional property emerges: *Orbital re-ordering impacts the sparsity of the CI Hamiltonian matrix and the number of non-vanishing CI coefficients in the CI eigensolutions*. This property is unique to CSF expansions and it is not present when a SD basis is utilized. This property follows from the way CSFs are constructed and coupled via the spin-free non-relativistic Hamiltonian operator, and it will be discussed in great detail in the present document. The wave function compression is here analyzed in combination with two orbital representations, commonly used in multi-configurational CI approaches, the *active natural orbitals* (NOs) produced by diagonalizing the active space one-body density matrix and, the *localized active orbitals* (LOs) that are obtained by localizing occupied and virtual orbital together (a not invariant transformation for HF wave functions). *Split-localized orbitals*, obtained by localizing occupied and virtual orbitals separately (an invariant transformation for HF wave functions) and mixed localized/delocalized basis represent alternative routes.

We would like to emphasize that, although we use different orbital representations, the aim of this work is not to compare between them, but rather to study the effect of re-ordering schemes on the ground state wave function sparsity in a spin-adapted basis within them. These re-ordering schemes are based on the occupation numbers for natural orbitals, real-space orbital separation arguments for localized orbitals, and generalized active space[48,52,54,58] orbital partitionings for both of them. Conventional CAS-CI procedures as well as the spin-adapted FCIQMC algorithm are used to show the wave function compression effect. The increased sparsity obtained for specific reordering schemes facili-

tates the convergence of spin-adapted FCIQMC calculations with respect to walker distributions, and it is strongly recommended for poly-nuclear transition metal complexes with antiferromagnetically coupled metal centers. The stretched N_2 and N_4 molecules, the chromium dimer and a model system of the oxidized form of the $[Fe_2S_2]$ cluster will be used as examples.

We discuss the theoretical foundation of the compression of spin-adapted wave functions in Section 2, and present numerical examples in Section 3, using conventional CI procedures and the stochastic FCIQMC algorithm. For the latter, we show that the convergence behavior with respect to the total number of walkers can be greatly improved by taking advantage of the wave function compression that follows orbital reordering.

2 Theoretical Details

2.1 Representation of CSFs

CSFs are generally represented by one of the three equivalent tables of Figure 1, known as Gel'fand, Paldus and Weyl tableaux, respectively.[43,59,60] The top row of the Gel'fand tableau

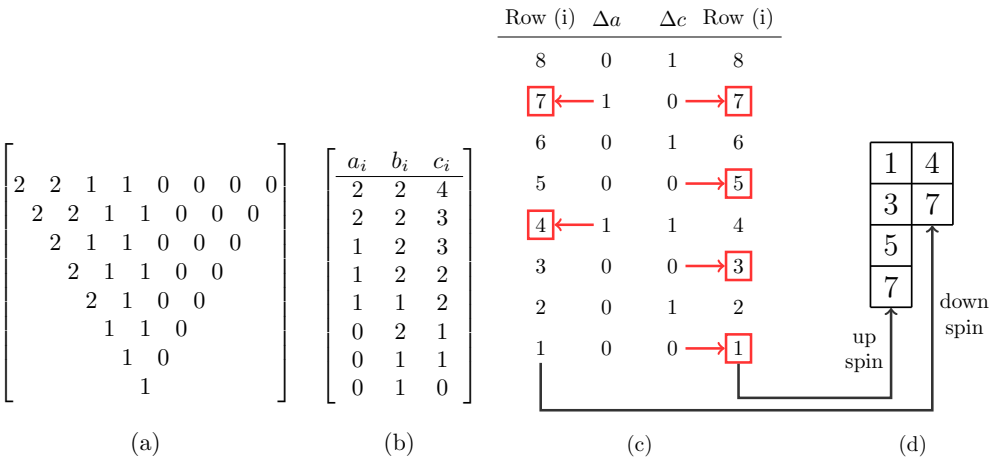

Figure 1: (a) Gel'fand, (b) Paldus ABC, (c) ΔAC variation and (d) Weyl tableau representing a distribution of 6 electrons in 8 orbitals with total spin $S = 1$.

completely characterizes the electronic state of the considered system as it complies with the

following two conditions

$$\sum_{i=1}^{n} m_{1i} = N, \quad \text{and} \quad \sum_{i=1}^{n} \delta_{1,m_{1i}} = 2S \tag{1}$$

where n, N and S are the total number of orbitals, electrons and the total spin, respectively. The m_{1i} elements represent the individual entries of the top-row. The example of Figure 1 represents a system with eight orbitals (n=8, dimension of the top-row) and six electrons (N=6, sum of the m_{1i} entries) coupled to a triplet spin state (S=1, sum of 1-entries divided by 2). The other rows in the Gel'fand tableau identify a specific CSF for the given electronic state, as it will be explained in the following. Considering that Gel'fand tableaux contain only 0, 1, and 2 entries,[43] a more compact "*three-column table*" can be used, where the number of 0's, 1's and 2's is counted. This table is referred to as the Paldus ABC tableau (Figure 1). The sum of the entries in Paldus ABC tableau equal the row index (from bottom to top),

$$a_i + b_i + c_i = i, \quad (i = 1, \ldots, n), \tag{2}$$

thus, any two columns are sufficient to uniquely determine the state, and the specific CSF. Paldus AC tableaux are derived from the ABC tableaux, by excluding the second column, B. Paldus ABC (or AC) tableaux can be recast in "*variation-tables*" with $\Delta x_i = x_i - x_{i-1}$ $(x = a, b, c)$, as shown in Figure 1. Starting from the top-row of Paldus ABC, or ΔABC variation tableau, four *actions* recursively follow to obtain the possible lower rows, and generate the CSFs for the targeted electronic state:

- remove one empty orbital, $\Delta a_i = 0$, $\Delta b_i = 0$, $\Delta c_i = 1$,
- reduce spin by $\frac{1}{2}$, $\Delta a_i = 0$, $\Delta b_i = 1$, $\Delta c_i = 0$ (negative spin-coupling),
- remove one doubly-occupied orbital and one empty orbital, and increase spin by $\frac{1}{2}$, $\Delta a_i = 1$, $\Delta b_i = -1$, $\Delta c_i = 1$ (positive spin-coupling), and
- remove one doubly-occupied orbital, $\Delta a_i = 1$, $\Delta b_i = 0$, $\Delta c_i = 0$

Lexically ordered CSFs are obtained when the steps above are followed in order. While the Δa and Δc entries are restricted to 0 and 1 values, the Δb column may assume 1, 0 and -1 entries. All CSFs for a given state can be constructed by allowing the possible variations of a_i, b_i and c_i according to the *actions* given above, decreasing the values of a_i, b_i and c_i down to (0 0 0). From the ΔAC tableaux, the Weyl representation is promptly obtained by writing the row-indices of the left 1-entries and the right 0-entries, as indicated in Figure 1. Each Weyl tableau represents a CSF with a defined total spin, S, and the left and right columns represent the positively and negatively spin coupled contributions in a cumulative sense.

The Step-Vector. The four possible actions that lead from the top-row of the ABC tableaux to the bottom can be expressed in a more compact form via the *step-vector*, defined as

$$d_i = 2\Delta a_i - \Delta c_i + 1. \tag{3}$$

Depending on the *action* that leads to the lower row index, the step-values will assume values from 0 to 3. Table 1 summarizes the correspondence between the possible step values and the Δa_i, Δb_i and Δc_i variations.

Table 1: Mapping between step-vector values, d_i, and the four possible variations of a_i, b_i, c_i and the equivalent nomenclature, d'_i, chosen in this manuscript.

d_i	Δa_i	Δb_i	Δc_i	d'_i
0	0	0	1	0
1	0	1	0	u
2	1	-1	1	d
3	1	0	0	2

Step-values, d_i, of 0, 1, 2 or 3 values correspond to empty, singly occupied orbitals increasing the total spin by 1/2 (positive spin-coupling and referred to as u in this work), singly occupied decreasing the total spin by 1/2 (negative spin-coupling and referred to as d) or doubly occupied ith-orbital, respectively. d'_i in Table 1 corresponds to the more intuitive step-value naming convention used to specify CSFs in the rest of this manuscript.

The Graphical Unitary Group Approach, GUGA. When constructing the CSFs of a given multi-configurational wave function, rows in Paldus ABC tableaux repeat for different CSFs. Repetitions can be avoided by listing only non-equivalent rows. The table collecting all the non-equivalent rows is referred to as a *distinct row table* (DRT) (Table 2), introduced by Shavitt.[39] Each row of a DRT is identified by a pair of indices, (i, j), with $i = a_i + b_i + c_i$ being the *level-index*, and j the *lexical row-index*, a counting index, such that $j < j'$ if $a_i > a'_i$ or if $a_i = a'_i$ and $b_i > b'_i$. CSFs are generated by connecting rows with decreasing level-index. Allowed connections between rows are indicated by *downward chaining indices*. For a given lexical row the downward chaining indices define the connected rows of the lower level row after the action of the four possible step-values, d_0, d_1, d_2 and d_3. Table 2 summarizes the DRT of a CAS(6,6) wave function, coupled to a singlet spin state. A more compact representation of DRT tables is obtained by means of *graphs* (Figure 2). Each vertex of the graph represents one distinct row of the DRT. Arcs connect only vertices linked by downward chaining indices. Vertices are labeled by the lexical ordering index, j, and arcs by the corresponding step-value. The *head-node* correspond to the top-row and the *tail-node* to the bottom-row (0 0 0) of the corresponding DRT table. Vertices with same i-value are aligned horizontally. Vertices are also left-right sorted with respect to the a and b values of the DRT. The left-right ordering ensures that the slope of each arc corresponds to its step value. *Direct walks* through the graph, following only vertices connected by arcs, lead to all possible CSFs of the given multi-configurational wave function. In Figure 2, three CSFs have been highlighted. The step-vector string associated to the orange path, $\mathbf{d} = |111222\rangle$, corresponds to the CSF $|uuuddd\rangle$ (u = positively spin-coupled, d = negatively spin-coupled), the green path to the CSF $|ududud\rangle$ and the blue path corresponds to the closed shell $|222000\rangle$ CSF (color online).

GUGA representation of GAS wave functions. When the GUGA representation of CSFs is used for *generalized active space* (GAS) wave functions,[48] a number of *direct walks* in the GUGA graph are not permitted by the *occupation number* constraints of the

Table 2: Distinct row table for N=6, n=6 and S=0. Zeroes under d_0–d_3 columns represent not allowed downward chaining. Paldus ABC representations of CSFs are obtained by selecting one row for each level index, i, according to the downwards chaining indices.

a	b	c	i	j	d_0	d_1	d_2	d_3
3	0	3	6	1	2	0	3	4
3	0	2	5	2	5	0	6	7
2	1	2	5	3	6	7	8	9
2	0	3	5	4	7	0	9	10
3	0	1	4	5	11	0	12	13
2	1	1	4	6	12	13	14	15
2	0	2	4	7	13	0	15	16
1	2	1	4	8	14	15	17	18
1	1	2	4	9	15	16	18	19
1	0	3	4	10	16	0	19	20
3	0	0	3	11	0	0	0	21
2	1	0	3	12	0	21	0	22
2	0	1	3	13	21	0	22	23
1	2	0	3	14	0	22	0	24
1	1	1	3	15	22	23	24	25
1	0	2	3	16	23	0	25	26
0	3	0	3	17	0	24	0	0
0	2	1	3	18	24	25	0	0
0	1	2	3	19	25	26	0	0
0	0	3	3	20	26	0	0	0
2	0	0	2	21	0	0	0	27
1	1	0	2	22	0	27	0	28
1	0	1	2	23	27	0	28	29
0	2	0	2	24	0	28	0	0
0	1	1	2	25	28	29	0	0
0	0	2	2	26	29	0	0	0
1	0	0	1	27	0	0	0	30
0	1	0	1	28	0	30	0	0
0	0	1	1	29	30	0	0	0
0	0	0	0	30	0	0	0	0

GAS specifications. A GAS6(6,6) is considered as an example, that contains six active electrons and six active orbitals, each orbital in a separate GAS sub-space. The six GAS sub-spaces are populated by only one electron and thus referred to as *disconnected spaces* (inter-space electron excitations are not allowed). This GAS wave function corresponds to a configurational space where only spin re-couplings via exchange-driven spin-flips are permitted. As a guide for the eye, the cumulative occupation number, N_{elec}, associated to each vertex is shown in Figure 2. Some of the arcs of Figure 2 are not permitted for this GAS6(6,6). For instance, the arc connecting vertices (30) and (27) corresponds to

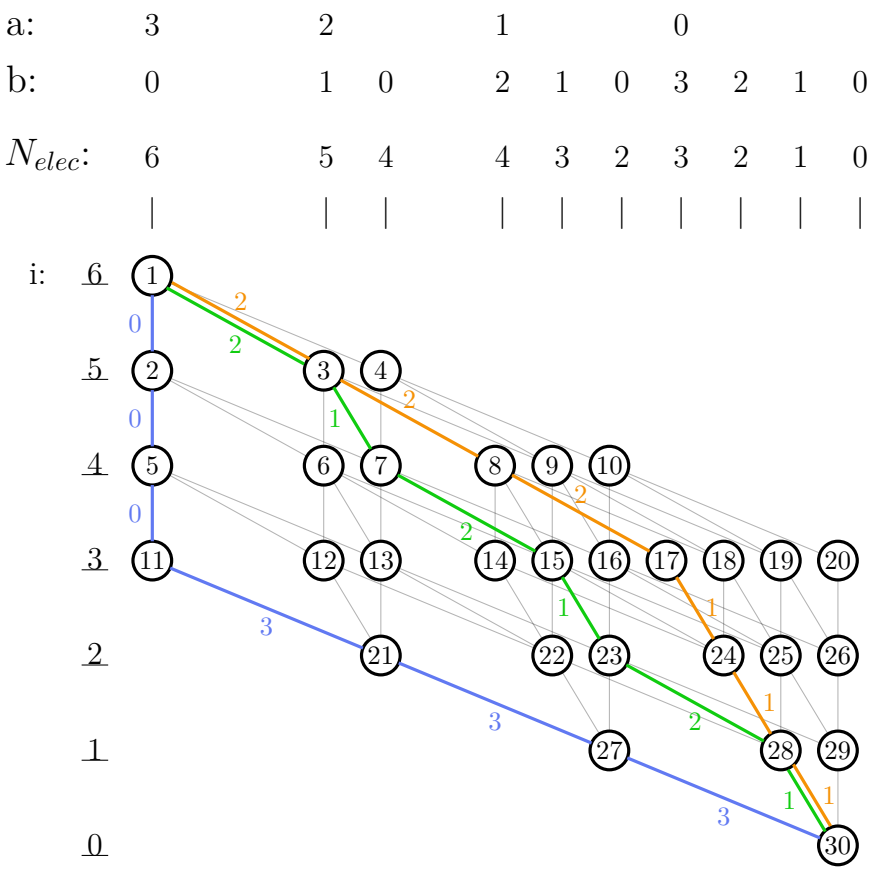

Figure 2: (Color online) Graph representing the DRT of Table 2. The different step-values, d_i, connecting the nodes of the highlighted paths are shown.

populating the first orbital with 2 electrons, which is not permitted by the chosen GAS. The paths permitted by the GAS6(6,6) restrictions are visualized in Figure 3. The filled black circles indicate the allowed vertices within the GAS restrictions. The thin black lines and circles of Figure 3, represent CSFs of the *auxiliary space*, a space that is forbidden by GAS rules, but necessary for the coupling of permitted CSFs via double excitations. The gray lines and circles are prohibited by GAS rules and not necessary for the auxiliary space.

2.2 Coupling of CSFs via the Hamiltonian operator

The coupling of two CSFs (or SDs) via the spin-free non-relativistic Hamiltonian operator

$$\hat{H} = \sum_{pq} h_{pq} \hat{E}_{pq} + \sum_{pq,rs} (pq|rs) \left(\hat{E}_{pq} \hat{E}_{rs} - \delta_{ps} \hat{E}_{rq} \right) = \sum_{pq} h_{pq} \hat{E}_{pq} + \sum_{pq,rs} (pq|rs) \, \hat{e}_{pq,rs} \qquad (4)$$

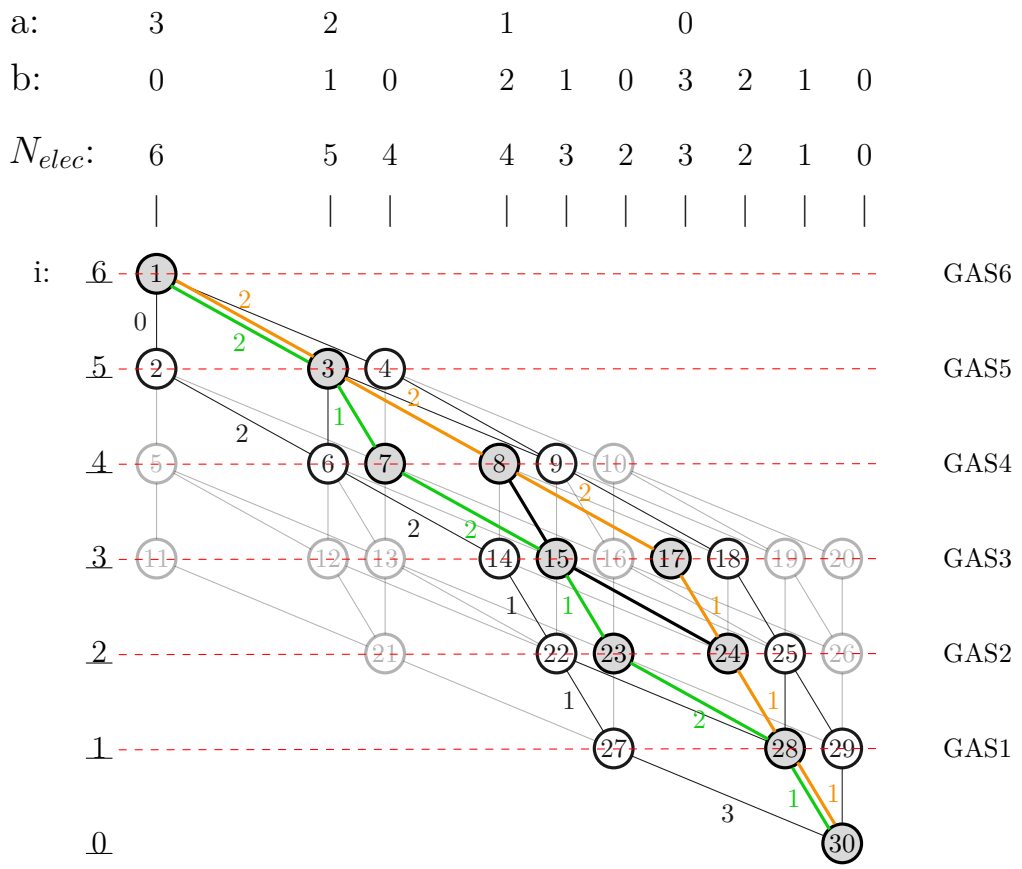

Figure 3: (Color online) Graph representing the DRT of Table 2 with the additional GAS6(6,6) constraints discussed in the main text. The orange, green and thick black paths between them represent the allowed CSFs of the GAS wave function, connecting the allowed vertices, indicated by the filled circles. The thin black lines and circles belong to the *auxiliary space* (see main text) and gray nodes and arcs are prohibited by GAS rules and not necessary for the auxiliary space.

is given by

$$\langle m'|\hat{H}|m\rangle = \sum_{pq} h_{pq}\langle m'|\hat{E}_{pq}|m\rangle + \sum_{pq,rs} (pq|rs)\,\langle m'|\hat{e}_{pq,rs}|m\rangle \tag{5}$$

where h_{pq} and $(pq|rs)$ are the one- and two-electron integrals and $\langle m'|\hat{E}_{pq}|m\rangle$ and $\langle m'|\hat{e}_{pq,rs}|m\rangle$ the *coupling coefficients* between two SDs or CSFs. Of course, the integrals depend on the *shape* of the orbitals, while the coupling coefficients depend on the entries in $|m\rangle$ and $|m'\rangle$ (depending if a SD or CSF basis is chosen). The Slater-Condon rules apply for the coupling coefficients between SDs, which can be evaluated very efficiently. However, these rules are not applicable for CSFs and consequently wave function optimizations in CSFs basis have

been less popular than optimizations in SDs basis. Paldus, Shavitt and others[40,41,47] have demonstrated that efficient evaluation of CSF coupling terms are possible via the GUGA approach.

One-electron coupling coefficients. One-electron coupling coefficients, $\langle m'|\hat{E}_{pq}|m\rangle$, can be computed graphically by first identifying $|m\rangle$ and $|m'\rangle$ paths in the corresponding GUGA graph followed by their connection via the excitation operator, $\hat{E}_{pq}$. For non-vanishing coefficients the walks of $|m'\rangle$ and $|m\rangle$ on the graph must coincide outside the (p,q) *range*, defined by the excitation operator, $\hat{E}_{pq}$. The value of the coupling coefficient is independent of the overlapping outer regions, and depends only on the shape of the *loop* formed by the two CSFs in the range defined by the operator $\hat{E}_{pq}$. At each row level, k, (inside the range) non-vanishing terms satisfy the conditions

$$\Delta b_k = b'_k - b_k = \pm 1, \tag{6}$$

$$2a'_k + b'_k = 2a_k + b_k + 1. \tag{7}$$

Shavitt proved that coupling coefficients can be factorized as

$$\langle m'|\hat{E}_{pq}|m\rangle = \prod_{k=p}^{q} W(Q_k; d'_k, d_k, \Delta b_k, b_k) \tag{8}$$

and values of $W(Q_k; d'_k, d_k, \Delta b_k, b_k)$ are tabulated.[61] The factors, W, depend on step vector values, d'_k and d_k, $\Delta b_k = b_k - b'_k$, and the b_k value of $|m\rangle$ at the k-level. They also depend on the *k-segment shape*, Q_k, which indicates the relation of the k-level arcs of the two CSFs, $|m'\rangle$ and $|m\rangle$. If the k-arc of $|m'\rangle$ is on the left, coincident or on the right of the k-arc of $|m\rangle$, Q_k is labeled as *raising* (R), *weight* (W) or *lowering* (L), respectively. For the segments where the loop begins (bottom) and ends (top) *under-bars* and *over-bars* ($\underline{R}, \underline{L}$ and $\overline{R}, \overline{L}$), respectively, are used as labels. As an example, the $\langle 2uud0d|\hat{E}_{15}|uuuddd\rangle$ term is promptly

evaluated using the labeling rules defined above, and Table III of Reference 61:

$$\langle 2uud0d|\hat{E}_{15}|uuuddd\rangle = \underline{R}^1_{31} \cdot {}^{+1}R^2_{11} \cdot {}^{+1}R^3_{11} \cdot {}^{+1}R^2_{22} \cdot \overline{R}^1_{22} \cdot W_{22} \tag{9}$$

where we have used the ${}^{\Delta b_k}R^{b_k}_{d'_k d_k}$ symbols for each k-level inside the loop.

When orbitals are re-ordered, the graphical representation of any CSF in a GUGA graph and the couplings between CSFs are altered and thus non-vanishing coupling terms may vanish after orbital re-ordering.

Two-electron coupling coefficients. Matrix elements of two-body excitation operators $\hat{E}_{pq}\hat{E}_{rs}$ can either be evaluated by introducing a summation over intermediate states, $|m''\rangle$ (resolution of identity),

$$\langle m'|\hat{E}_{pq}\hat{E}_{rs}|m\rangle = \sum_{m''} \langle m'|\hat{E}_{pq}|m''\rangle\langle m''|\hat{E}_{rs}|m\rangle. \tag{10}$$

or directly in factorized form similar to Eq. (8), see Reference 40. Similar to the one-body coupling coefficients, orbital re-ordering also impacts these terms, thus, it is possible to increase the number of vanishing coupling terms and produce a more sparse Hamiltonian matrix and compact representation of the many-body wave function.

3 Applications

This section is dedicated to examples that show how sparsity of spin adapted CI wave functions is increased by orbital reordering and rotations.

3.1 The Nitrogen Molecule

The nitrogen molecule at dissociation is used to demonstrate how orbital re-ordering impacts the sparsity of the many-body CI expansion when CSF representations are utilized. Consider a CAS(6,6) active space, consisting of six MOs formed by linear combination of the $2p$ atomic

orbitals on each atom, and their electrons. The wave function is optimized to a singlet spin state. Two set of orbitals are considered: delocalized *natural orbitals*, with bonding and anti-bonding character, and *localized orbitals* (atomic-orbital-like). C_1 point group symmetry has been used for all cases. The CAS(6,6) CI expansion contains 175 CSFs. These CSFs are represented by all possible *walks* in the GUGA graph of Figure 2. The number and the list of non-vanishing terms in the optimized CI wave function for each type of orbital shape and ordering are given in Table 4 and Table 5.

Natural orbitals. Two ordering schemes have been adopted for the natural orbitals of the CAS(6,6) wave function, the *canonical ordering*, with bonding orbitals (σ, π_x and π_y) preceding the anti-bonding orbitals (σ^*, π_x^* and π_y^*), and the *pair ordering*, where orbitals are sorted in (σ, σ^*), (π_x, π_x^*) and (π_y, π_y^*) pairs. The *pair ordering* has the effect of reducing the number of non-vanishing terms in the CI expansion with respect to the *canonical ordering*, from 20 to 14 CSFs (Table 4).

In the NO basis there is a strong coupling between bonding and anti-bonding orbital pairs, i.e. $\sigma \leftrightarrow \sigma^*$ and $\pi_{x/y} \leftrightarrow \pi_{x/y}^*$. As a consequence, the significant off-diagonal molecular integrals at dissociation are the exchange-like, $(\sigma\sigma^*|\sigma\sigma^*)$ and $(\pi_{x/y}\pi_{x/y}^*|\pi_{x/y}\pi_{x/y}^*)$, and the "coherent" combinations of them, i.e. $(\sigma\sigma^*|\pi_{x/y}\pi_{x/y}^*)$ (8-fold permutational symmetry implied). We will explain the increased sparsity of the pair ordering scheme using the example of three CSFs shown in Table 3.

Table 3: Three exemplary CSFs of the CAS(6,6) of N_2 at dissociation in a NO basis.

Canonical order:	σ	π_x	π_y	π_y^*	π_x^*	σ^*
$\|1\rangle$:	2	2	2	0	0	0
$\|2\rangle$:	2	u	d	u	d	0
$\|3\rangle$:	2	u	u	d	d	0
Pair order:	σ	σ^*	π_x	π_x^*	π_y	π_y^*
$\|1'\rangle$:	2	0	2	0	2	0
$\|2'\rangle$:	2	0	u	d	u	d
$\|3'\rangle$:	2	0	u	u	d	d

The coupling between states $|1\rangle \leftrightarrow |2\rangle$ and $|1\rangle \leftrightarrow |3\rangle$ in both ordering schemes is driven by the large integral contributions $(\pi_x\pi_x^*|\pi_y\pi_y^*)$. In the canonical ordering the coupling coefficient is non-zero for both $\langle 2|\hat{e}_{52;43}|1\rangle$ and $\langle 3|\hat{e}_{52;43}|1\rangle$. The coupling between states $|2\rangle$ and $|3\rangle$ is driven by the exchange contributions $(\pi_{x/y}\pi_{x/y}^*|\pi_{x/y}\pi_{x/y}^*)$ and the Hamiltonian matrix element is given by

$$\langle 2uudd0|\hat{H}|2udud0\rangle = \frac{\sqrt{3}}{2}\left[(34|34) + (25|25) - (35|35) - (24|24)\right],$$

where $(34|34)$ and $(25|25)$ correspond to the large and identical integrals between the bonding and anti-bonding π orbitals. Thus in the canonical ordering configurations $|1\rangle$, $|2\rangle$ and $|3\rangle$ are coupled, causing the wave function to be dense.

The coupling coefficient between state $|1'\rangle$ and $|3'\rangle$ for the pair ordering is zero for the strong $(\pi_x\pi_x^*|\pi_y\pi_y^*)$ integral contribution and is only driven by the much smaller $(\pi_x\pi_y|\pi_x^*\pi_y^*)$ and similar contributions. At the same time, in the pair ordered case, the matrix element

$$\langle 20uudd|\hat{H}|20udud\rangle = \frac{\sqrt{3}}{2}\left[(45|54) + (36|63) - (36|64) - (35|53)\right],$$

cancels to zero, as all the involved integrals corresponds to identical—and weak—$(\pi_x\pi_y|\pi_x\pi_y)$ types.

To summarize, the pair ordering reduces the connectivity within the Hilbert space, by either vanishing coupling coefficients for strong integral contributions or cancellation of equal integral contribution, due to the sign structure of the resulting coupling coefficients. This reduced connectivity within CSFs, leads to a more sparse ground state wave function in a NO basis with pair ordering.

Localized orbitals. Localization of the natural orbitals produces atomic-orbital like molecular orbitals. The *pair ordering* scheme $\left(p_x^Ap_x^B, p_y^Ap_y^B, p_z^Ap_z^B\right)$ and the *atom separated ordering* scheme $\left(p_x^Ap_y^Ap_z^A, p_x^Bp_y^Bp_z^B\right)$, have been considered.

In the *pair ordering* scheme the wave function contains 5 non-vanishing terms. A single-

configurational (yet multi-determinantal) wave function is obtained when the *atom separated ordering* is adopted. Localization schemes without particular attention on the orbital ordering utilized is not a sufficient condition for optimal wave function compression.

The five non-vanishing CSFs, for the *pair ordered* localized orbitals, are graphically shown in Figure 2. They are all the ones inside (and including) the orange and green *direct walks*. These CSFs share a common property: they all feature singly-occupied orbitals. Thus, a GAS wave function can be constructed, with each orbital in a separate GAS sub-space, and the spaces kept *disconnected*. The graph of Figure 3 represents such a wave function.

Table 4: Number of non-vanishing CSFs in the CAS(6,6) of N_2 and the CAS(12,12) of N_4 at dissociation geometry.

Shape	Ordering	N_2 system	N_4 system
Delocalized	Canonical	20	2073[(a)]
Delocalized	Pair/Type	14	1100[(a)]
Localized	Pair/Type	5	119
Localized	Atom separated	1	20

(a) These values may change as a function of the local rotations of the p_x and p_y orbitals at each site. See main text for details.

The single configurational character of the CI expansion in the localized orbitals and *atom-separated ordering* completely reflects the chemical nature of this system, that is, two non-interacting nitrogen atoms, each in its ground state, ^4S, anti-ferromagnetically coupled to form a singlet spin-state compound. This information is not promptly accessible when orbitals are delocalized or localized and *pair ordered*.

The GAS Hamiltonian matrix. The single-configurational character of the wave function in *atom separated ordering* scheme is bound to the sparsity of the corresponding Hamiltonian matrix. Only a matrix with vanishing off-diagonal elements can provide a strictly single-configurational eigenvector. For simplicity and without loss of generality, only the GAS6(6,6) Hamiltonian matrix is discussed. The GAS6(6,6) wave function contains a total of five CSFs (listed in the third column of Table 5). The GAS6(6,6) Hamiltonian matrices in the localized orbital basis and using *pair ordering* and *atom separated ordering* are reported in Figure 4.

Table 5: List of non-vanishing CSFs for the CAS(6,6) wave function of N_2 at dissociation. Natural orbitals and localized orbitals are shown. Natural orbitals in *canonical ordering* are sorted as $\left(\sigma\pi_x\pi_y, \sigma^*\pi_x^*\pi_y^*\right)$. Natural orbitals in *pair ordering* are sorted as $\left(\sigma\sigma^*, \pi_x\pi_x^*, \pi_y\pi_y^*\right)$. Localized orbitals in *pair ordering* are sorted as $\left(p_x^A p_x^B, p_y^A p_y^B, p_z^A p_z^B\right)$. Localized orbitals in *atom separated ordering* are sorted as $\left(p_x^A p_y^A p_z^A, p_x^B p_y^B p_z^B\right)$.

Natural Orbitals		Localized Orbitals	
Canonical Ordering	Pair Ordering	Pair Ordering	Atom Separated
222000	202020	ududud	uuuddd
220200	022020	uduudd	
2udud0	200220	uuddud	
u2du0d	020220	uududd	
202020	202002	uuuddd	
ud20ud	022002		
022002	200202		
2uudd0	020202		
u2ud0d	uudd20		
uu20dd	uu20dd		
200220	20uudd		
ud02ud	02uudd		
020202	uu02dd		
u0du2d	uudd02		
0udud2			
002022			
uu02dd			
u0ud2d			
0uudd2			
000222			

$$
\begin{array}{c}
\begin{array}{ccccc} \text{uuuddd} & \text{uududd} & \text{uduudd} & \text{uuddud} & \text{ududud} \end{array} \\
\begin{pmatrix}
-107.22 & & & & \\
-0.02 & -107.21 & & & \\
0.03 & -0.02 & -107.23 & & \\
0.03 & -0.02 & 0.04 & -107.23 & \\
-0.06 & 0.04 & -0.07 & -0.07 & -107.31
\end{pmatrix}
\end{array}
\qquad
\begin{array}{c}
\begin{array}{ccccc} \text{uuuddd} & \text{uududd} & \text{uduudd} & \text{uuddud} & \text{ududud} \end{array} \\
\begin{pmatrix}
-107.44 & & & & \\
0.00 & -107.19 & & & \\
0.00 & 0.00 & -107.19 & & \\
0.00 & 0.00 & 0.00 & -107.19 & \\
0.00 & 0.00 & 0.00 & 0.00 & -107.19
\end{pmatrix}
\end{array}
$$

Figure 4: GAS6(6,6) Hamiltonian matrices on the basis of localized orbitals in *pair ordering* (left) and *atom separated ordering* (right).

As an example, two off-diagonal elements, $\langle ududud|\hat{H}|uuuddd\rangle$ and $\langle uduudd|\hat{H}|uuuddd\rangle$ are evaluated, and it is shown why in the *atom separated ordering* these terms vanish while in the *pair ordering* they do not.

General one-electron excitation operators, $\hat{E}_{pq}(p \neq q)$ applied to any of the CSFs of the GAS6(6,6) wave function necessarily generate CSFs outside the GAS expansion (CSFs with doubly occupied orbitals), and no contribution to the off-diagonal elements of the GAS Hamiltonian matrix can arise from them. Only *exchange two-particle operators* can contribute to these elements (double spin-flips), namely

$$\langle m'|\hat{H}_{GAS}|m\rangle = \frac{1}{2}\sum_{pq}(pq|qp)\,\langle m'|\hat{E}_{pq}\hat{E}_{qp}|m\rangle. \tag{11}$$

The resolution-of-identity (Equation 10) is used for their evaluation. The $|m''\rangle$ configurations of the *auxiliary space*, that simultaneously couple with $|uuuddd\rangle$ and $|ududud\rangle$ (or $|uduudd\rangle$), are found by applying conditions (6) and (7). The black thin lines of Figure 3 represent the CSFs of the auxiliary space that simultaneously couple with $|uuuddd\rangle$ and $|ududud\rangle$, and the resulting coupling coefficients are listed in Table 6.

Table 6: Non-vanishing $(pq|qp)\,\langle ududud|\hat{E}_{pq}\hat{E}_{qp}|uuuddd\rangle$ terms.

Term	Value
$(26\|62)\,\langle ududud\|\hat{E}_{62}\|u2udd0\rangle\langle u2udd0\|\hat{E}_{26}\|uuuddd\rangle$	$-\frac{\sqrt{2}}{2}\,(26\|62)$
$(15\|51)\,\langle ududud\|\hat{E}_{51}\|2uud0d\rangle\langle 2uud0d\|\hat{E}_{51}\|uuuddd\rangle$	$-\frac{\sqrt{2}}{2}\,(15\|51)$
$(26\|62)\,\langle ududud\|\hat{E}_{26}\|u0udd2\rangle\langle u0udd2\|\hat{E}_{62}\|uuuddd\rangle$	$-\frac{\sqrt{2}}{2}\,(26\|62)$
$(15\|51)\,\langle ududud\|\hat{E}_{15}\|0uud2d\rangle\langle 0uud2d\|\hat{E}_{51}\|uuuddd\rangle$	$-\frac{\sqrt{2}}{2}\,(15\|51)$

The coupling terms with the auxiliary $|u2ud0d\rangle$, $|2uudd0\rangle$, $|0uudd2\rangle$ and $|u0ud2d\rangle$ CSFs vanish as the corresponding (25|52) and (16|61) integrals, that are multiplied with, equal zero in both orbital representations. As a result

$$\langle ududud|\hat{H}|uuuddd\rangle = \frac{\sqrt{2}}{2}\left[(26|62) + (15|51)\right], \tag{12}$$

and

$$\langle uduudd|\hat{H}|uuuddd\rangle = \frac{\sqrt{2}}{3}\left[(26|62) + (15|51)\right] + \frac{2\sqrt{2}}{3}\left[(24|42) + (35|53)\right]. \tag{13}$$

For the *atom separated ordering* case, the two-electron repulsion integrals of Equation (12) and (13) vanish, as the orbitals of the pairs $(2,6)$, $(1,5)$, $(2,4)$ and $(3,5)$ are spatially separated. In the *pair ordering* case, instead the orbitals of these pairs reside on same atom and the two-electron repulsion integrals do not vanish, leading to a non-vanishing Hamiltonian matrix element.

3.2 Square N_4 System

In this section the square N_4 model compound at dissociation and in its singlet spin-state is discussed. A CAS(12,12) active space that consists of the three $2p$ orbitals on each atom and their electrons is chosen. The corresponding CI expansion contains a total of 226512 CSFs in the C_1 point group symmetry. Natural and localized orbitals are used as basis for the CI procedure, using *canonical*, *type*, and *atom separated* orderings. The *type ordering* for the localized orbitals is the following:

$$(2p_x^A 2p_x^B 2p_x^C 2p_x^D)(2p_y^A 2p_y^B 2p_y^C 2p_y^D)(2p_z^A 2p_z^B 2p_z^C 2p_z^D).$$

The number of non-vanishing terms for each orbital representation is summarized in Table 4. As for the nitrogen molecule, the most compact representation of the CI wave function is obtained when localized orbitals in *atom separated* ordering are utilized. The corresponding wave function contains only 20 non-vanishing terms. This number can be derived from a GAS12(12,12) wave function with disconnected spaces and singly occupied orbitals only. Under these conditions, the first three electrons, residing on the first atom, are coupled to a quartet, the last three electrons are coupled anti-ferromagnetically to the previous ones, locally with parallel spins, while the intermediate six electrons couple in all possible ways with parallel or anti-parallel spin. The resulting CSFs are represented graphically in the *genealogical branching diagram* of Figure 5. In the *Type ordering* the gray paths of Figure 5b

do not vanish and, instead, they contribute to the CI expansion.

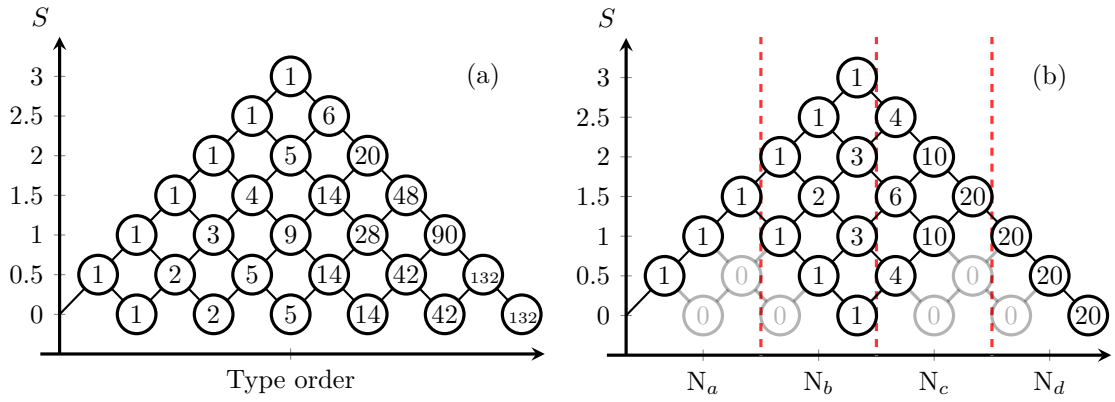

Figure 5: Genealogical branching diagram for the N_4 model system at dissociation in the localized orbital basis and considering only spin-flips of singly occupied orbitals with type ordering (a) and atom separated ordering (b).

3.3 The Chromium dimer, CAS(24,48).

Small active spaces have been considered for the N_2 and N_4 systems, with wave functions that can be optimized by conventional methods (Davidson in its Direct-CI formalism). GAS restrictions, similar to the nitrogen cases, discussed in Sections 3.1 and 3.2, could also be applied to the valence orbitals of the chromium dimer at dissociation in a localized basis. Consequently the valence only active space for the chromium dimer at dissociation, CAS(12,12), can be related to the CAS(6,6) of the nitrogen molecule, with the two chromium atoms in their high-spin, ^7S, ground state and anti-ferromagnetically coupled. This CAS(12,12) wave function would be single configurational if represented by localized orbitals in *atom separated* ordering, and will not be discussed further.

Instead, we are interested in a considerably larger active space, CAS(24,48) (see Reference 17 for details on the active space). Conventional CI optimization procedures are prohibitive, and the spin-adapted implementation of the FCIQMC algorithm has been utilized. In the CAS(24,28) case, doubly occupied, singly occupied and empty orbitals simultaneously

occur in the wave function, and a mixture of static and dynamic correlation contributions characterizes this wave function, independently of the orbital representation chosen. Two geometries are discussed, one at the dissociation limit and one at a bond distance of 2.4Å—the "shoulder region" of the potential energy curve—where a more complex wave function is to be expected.

Dissociation limit. Figure 6 shows four spin-adapted FCIQMC calculations using (a) delocalized orbitals in *canonical ordering*, (b) delocalized orbitals in *pair ordering*, (c) localized orbitals in *pair ordering*, and (d) localized orbitals in *atom separated ordering*. For case (b), pair ordering was adopted for all orbitals and, orbitals have been reordered such that orbitals with p and s character (any shell) are in adjacent positions. In case (d), orbitals have been sorted as

$$\left(3p4p5p4f4d5s3d4s\right)_A \left(4s3d5s4d4f5p4p3p\right)_B \tag{14}$$

An unstable spin-adapted FCIQMC dynamics is observed for the delocalized orbitals in canonical ordering even at a population of 20×10^6 walkers (20M). The wave function is highly multi-configurational in this orbital representation, and walkers are evenly distributed among the many equivalent configurations. As a consequence the occupation of the reference CSF drops almost to zero, which in turn caused the energy estimate to diverge. It has been observed already for the simpler N_2 and N_4 cases that *pair ordering* reduces the number of non-vanishing configurations. In the present case this reduction is sufficient to stabilize the FCIQMC dynamics. Yet, with a population of 20×10^6 walkers the energy is not converged.

The localization schemes improve the dynamics, and the *atom separated reordering* has a major effect on the sparsity of the wave function. Already with a population of 1×10^6 walkers (1M) a satisfactory dynamics is observed with a projected energy estimate ~ 30 kcal/mol lower than the case with localized orbitals in *pair ordering*, and ~ 40 kcal/mol lower than the energy estimate obtained with delocalized orbitals and with higher walker population

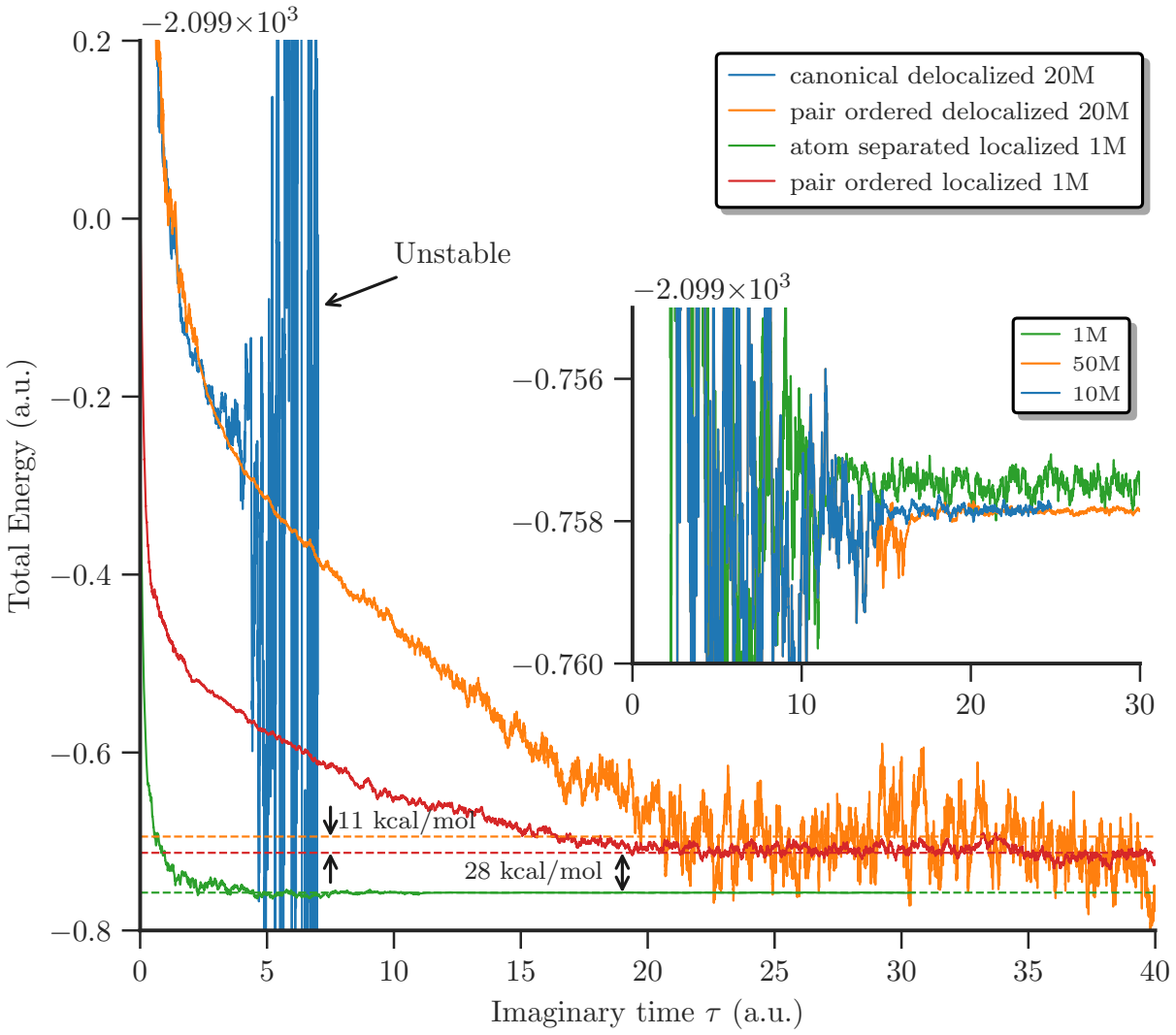

Figure 6: Spin adapted FCIQMC dynamics for the chromium dimer at the dissociation limit, using a (24,48) active space.

(20M). Furthermore, increasing the walker population from 1×10^6 to 50×10^6 walkers causes a marginal lowering of the projected energy, by less than 1 milliHartree, indicating that convergence with respect to walker population has been reached.

Intermediate bond distance. Also at the intermediate bond distance of 2.4 Å the orbital representation has an impact on the wave function sparsity (Figure 7). An unstable dynamics is observed also in this case, when delocalized orbitals in canonical ordering are used. The dynamics improves when orbitals are ordered in bonding and anti-bonding pairs

and even more when localized orbitals in *atom separated* ordering are utilized. The latter representation lead to a FCIQMC dynamics that, already at 5×10^6 walker population, is 22 kcal/mol lower than the dynamics on the basis of the delocalized and *pair ordered* orbitals and higher walker population (20M).

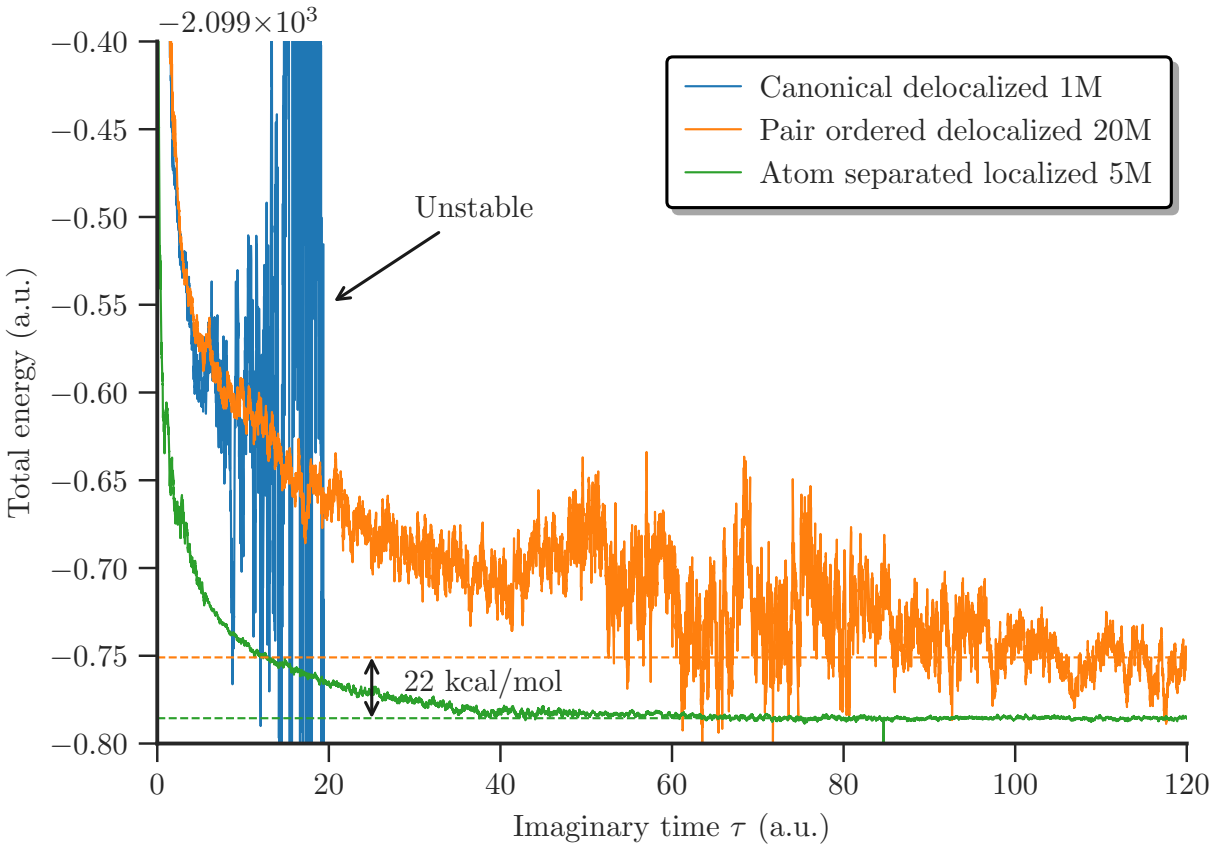

Figure 7: Spin adapted FCIQMC dynamics for the chromium dimer at a bond distance of 2.4Å, using the (24,48) active space.

3.4 The Fe_2S_2 model system.

In this section we show that orbital reordering can lead to higher sparsity in spin adapted CI wave functions also in practical cases, where no obvious simplifications of the wave function can easily be predicted. An active space of 22 electrons and 26 orbitals is considered for a $[Fe_2^{(III)}S_2]^{2-}$ model system (coordinates available in the SI),[62] that consists of the 20 valence,

$3d$, and double-shell, d', orbitals on the metal centers, and the 6 $3p$ orbitals of the bridging sulfur atoms. Formally the sulfur orbitals are doubly occupied (12 electrons, S^{2-}), and the iron atoms are in their Fe(III) oxidation state (d^5 configuration, 10 electrons), for a total of 22 electrons. The low-spin state (singlet) with anti-ferromagnetically coupled spins at the metal centers is characterized by a highly correlated wave function (details of the wave function go beyond the scope of the present work).

Spin-adapted FCIQMC wave function optimizations have been performed on the basis of the Stochastic-CASSCF(22,26)[36] optimized natural orbitals. Two orbital ordering schemes are discussed for this system, the *canonical ordering* and the *pair ordering*.
Similar to the Cr_2 case, the spin-adapted FCIQMC dynamics are highly unstable, when delocalized orbitals in *canonical ordering* are utilized (see Figure 8). The dynamics becomes stable after the natural orbitals have been reordered following a qualitative *pair ordering* scheme (yellow line in Figure 8). The pair reordering in this case is *qualitative* due to the mixing of the sulfur atomic orbitals into the metal centered molecular orbitals. The horizontal green line in Figure 8 corresponds to the Stochastic-CASSCF(22,26) energy in a SD basis and using 1×10^8 walkers (100M). The spin-adapted FCIQMC energy estimate using the same number of walkers is only ~ 3 milliHartree above. This example shows the paramount importance of a careful orbital re-ordering to enable stable FCIQMC dynamics.

4 Conclusions and Outlook

We have demonstrated that the sparsity of multi-configurational wave functions expanded in CSFs depends on the orbital ordering, as well as orbital representation, the former feature being unique to CSF expansions. Orbital transformations can be applied that greatly reduce the number of non-vanishing CI coefficients of multiconfigurational wave functions. Bonding and anti-bonding pair ordering for delocalized orbitals, and atom separated ordering schemes for localized orbitals maximally increase the sparsity of spin-adapted wave functions for

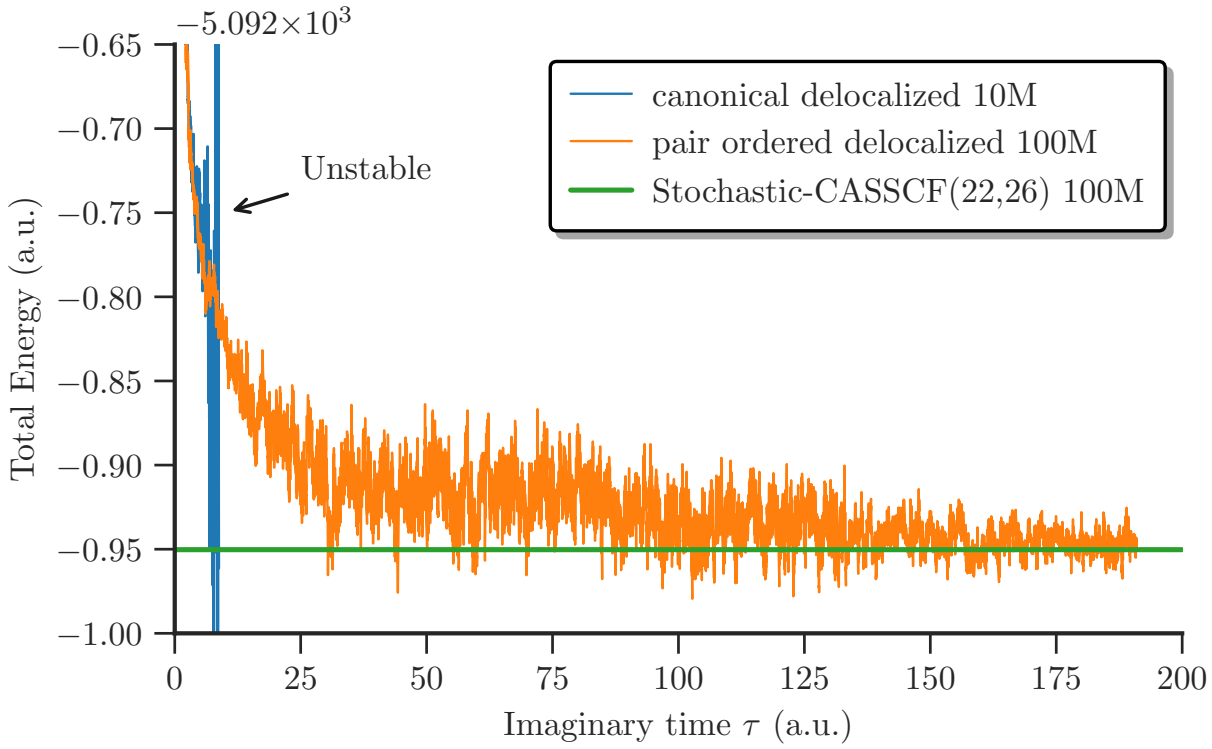

Figure 8: Spin adapted FCIQMC dynamics for the $[Fe_2S_2]^{2-}$ cluster, using a CAS(22,26) active space. The green line corresponds to the converged Stochastic-CASSCF(22,26) wave function using 1×10^8 walkers, and lies at -5092.9503 a.u. By increasing the total number of walkers to 2×10^9, the Stochastic-CASSCF calculation in a SD basis converges to a total energy of -5092.9512 a.u., that can be used for future references.

the examples discussed in this work. The increased sparsity that follows from the orbital reordering is greatly beneficial for methods that approximate FCI wave functions, such as FCIQMC. This approach greatly enhances the convergence of the spin adapted FCIQMC algorithm, and in certain difficult cases is found to be the only viable way to stable dynamics. The protocol is general and can be applied to molecular systems of practical interest. A simple reordering of the CAS(22,26) natural orbitals of a $[Fe_2S_2]^{2-}$ model system has been discussed.

The aim of this work is to show the effect of re-ordering schemes *within* a chosen orbital representation (delocalized or localized), and not to compare between them. However, our results indicate that for complex molecular systems a mixed delocalized / localized orbital representation is preferable, using a delocalized and *pair ordered* representation for orbitals

that describe *covalent bonds*, and a localized and *atom separated* representation for singly occupied orbitals (which will be investigated more thoroughly in future work). Localization of covalent bonds can easily lead to unnecessary complications at the level of the wave function, with new terms appearing to account for the orbital mixing. On the other hand a localized representation for singly occupied orbitals guarantees the highest sparsity for anti-ferromagnetically coupled polynuclear spin systems, removing the unnecessary entanglement of electrons that follows from using delocalized orbitals for unpaired electrons (as already shown for the simple N_2 system).

5 Supplementary materials

The Supplementary Material contains details on: Listing 1: Cartesian Coordinates for the FeS cluster model. Listint 2: OpenMolcas input for the FeS cluster model. Listint 3: NECI input in SD basis for the FeS cluster model. Listint 4: NECI input in CSF basis for the FeS cluster model. Listint 5: OpenMolcas input for the chromium dimer.

6 Competing Financial Interests statement

The author declares no competing financial interest.

7 Acknowledgments

GLM thanks Quan Phung for inspiring discussions about the Fe_2S_2 model system.

8 Correspondence

Correspondence and requests for materials should be addressed to `g.limanni@fkf.mpg.de` or `giovannilimanni@gmail.com`.

References

(1) Bogdanov, N. A.; Li Manni, G.; Sharma, S.; Gunnarsson, O.; Alavi, A. New superexchange paths due to breathing-enhanced hopping in corner-sharing cuprates. *Arxiv* **2018**,

(2) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. STRUCTURE, FUNCTION, AND FORMATION OF BIOLOGICAL IRON-SULFUR CLUSTERS. *Annual Review of Biochemistry* **2005**, *74*, 247–281.

(3) Noodleman, L.; Peng, C.; Case, D.; Mouesca, J.-M. Orbital interactions, electron delocalization and spin coupling in iron-sulfur clusters. *Coordination Chemistry Reviews* **1995**, *144*, 199 – 244.

(4) Zener, C. Interaction Between the *d* Shells in the Transition Metals. *Phys. Rev.* **1951**, *81*, 440–444.

(5) Anderson, P. W.; Hasegawa, H. Considerations on Double Exchange. *Phys. Rev.* **1955**, *100*, 675–681.

(6) Girerd, J. Electron transfer between magnetic ions in mixed valence binuclear systems. *The Journal of Chemical Physics* **1983**, *79*, 1766–1775.

(7) Noodleman, L.; Baerends, E. J. Electronic structure, magnetic properties, ESR, and optical spectra for 2-iron ferredoxin models by LCAO-X.alpha. valence bond theory. *Journal of the American Chemical Society* **1984**, *106*, 2316–2327.

(8) Noodleman, L.; Davidson, E. R. Ligand spin polarization and antiferromagnetic coupling in transition metal dimers. *Chemical Physics* **1986**, *109*, 131 – 143.

(9) Gibson, J. F.; Hall, D. O.; Thornley, J. H.; Whatley, F. R. The iron complex in spinach ferredoxin. *Proceedings of the National Academy of Sciences* **1966**, *56*, 987–990.

(10) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A Complete Active Space SCF Method (CASSCF) Using a Density Matrix Formulated Super-CI Approach. *Chem. Phys.* **1980**, *48*, 157–173.

(11) Roos, B. O. The Complete Active Space SCF Method in a Fock-Matrix-Based Super-CI Formulation. *Int. J. Quantum Chem.* **1980**, *18*, 175–189.

(12) Siegbahn, P. E. M.; Heiberg, A.; Roos, B. O.; Levy, B. A Comparison of the Super-CI and the Newton-Raphson Scheme in the Complete Active Space SCF Method. *Phys. Scr.* **1980**, *21*, 323–327.

(13) Siegbahn, P. E. M.; Almlöf, J.; Heiberg, A.; Roos, B. O. The Complete Active Space SCF (CASSCF) Method in a Newton–Raphson Formulation with Application to the HNO Molecule. *J. Chem. Phys.* **1981**, *74*, 2384–2396.

(14) Roos, B. O. *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2, Volume 69*; 2007; Vol. 69; pp 399–445.

(15) La Macchia, G.; Li Manni, G.; Todorova, T. K.; Brynda, M.; Aquilante, F.; Roos, B. O.; Gagliardi, L. On the Analysis of the Cr-Cr Multiple Bond in Several Classes of Dichromium Compounds. *Inorganic Chemistry* **2010**, *49*, 5216–5222.

(16) Li Manni, G.; Dzubak, A. L.; Mulla, A.; Brogden, D. W.; Berry, J. F.; Gagliardi, L. Assessing Metal-Metal Multiple Bonds in Cr-Cr, Mo-Mo, and W-W Compounds and a Hypothetical U-U Compound: A Quantum Chemical Study Comparing DFT and Multireference Methods. *Chemistry: A European Journal* **2012**, *18*, 1737–1749.

(17) Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS Method for Strong Correlation and the Challenging Case of Cr2. *J. Chem. Theory Comput.* **2013**, *9*, 3375–3384.

(18) Brogden, D. W.; Turov, Y.; Nippe, M.; Li Manni, G.; Hillard, E. A.; ClÃČÂľrac, R.; Gagliardi, L.; Berry, J. F. Oxidative Stretching of Metal-Metal Bonds to Their Limits. *Inorganic Chemistry* **2014**, *53*, 4777–4790.

(19) White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **1992**, *69*, 2863–2866.

(20) White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B* **1993**, *48*, 10345–10356.

(21) Zgid, D.; Nooijen, M. The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space. *The Journal of Chemical Physics* **2008**, *128*, 144116.

(22) Ghosh, D.; Hachmann, J.; Yanai, T.; Chan, G. K.-L. Orbital optimization in the density matrix renormalization group, with applications to polyenes and Îš-carotene. *The Journal of Chemical Physics* **2008**, *128*, 144117.

(23) Keller, S.; Dolfi, M.; Troyer, M.; Reiher, M. *J. Chem. Phys.* **2015**, *143*, 244118.

(24) Kurashige, Y.; Yanai, T. *J. Chem. Phys.* **2009**, *130*, 234114.

(25) Marti, K. H.; Ondík, I. M.; Moritz, G.; Reiher, M. Density Matrix Renormalization Group Calculations on Relative Energies of Transition Metal Complexes and Clusters. *J. Chem. Phys.* **2008**, *128*, 014104.

(26) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. *Nat. Chem.* **2014**, *6*, 927–933.

(27) Kurashige, Y. Multireference Electron Correlation Methods with Density Matrix Renormalisation Group Reference Functions. *Mol. Phys.* **2014**, *112*, 1485–1494.

(28) Kurashige, Y.; Chan, G. K.-L.; Yanai, T. *Nat. Chem.* **2013**, *5*, 660–666.

(29) Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *The Journal of Chemical Physics* **2009**, *131*, 054106.

(30) Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *The Journal of Chemical Physics* **2010**, *132*, 041103.

(31) Cleland, D.; Booth, G. H.; Alavi, A. A Study of Electron Affinites Using the Initiator Approach to Full Configuration Interaction Quantum Monte Carlo. *J. Chem. Phys.* **2011**, *134*, 024112.

(32) Overy, C.; Booth, G. H.; Blunt, N. S.; Shepherd, J. J.; Cleland, D.; Alavi, A. Unbiased Reduced Density Matrices and Electronic Properties from Full Configuration Interaction Quantum Monte Carlo. *J. Chem. Phys.* **2014**, *141*, 244117.

(33) Blunt, N. S.; Smart, S. D.; Kersten, J. A.-F.; Spencer, J. S.; Booth, G. H.; Alavi, A. Semi-Stochastic Full Configuration Interaction Quantum Monte Carlo: Developments and Application. *J. Chem. Phys.* **2015**, *142*, 184107.

(34) Booth, G. H.; Smart, S. D.; Alavi, A. Linear-Scaling and Parallelisable Algorithms for Stochastic Quantum Chemistry. *Mol. Phys.* **2014**, *112*, 1855–1869.

(35) Blunt, N. S.; Alavi, A.; Booth, G. H. *Phys. Rev. Lett.* **2015**, *115*, 050603.

(36) Li Manni, G.; Smart, S. D.; Alavi, A. Combining the Complete Active Space Self-Consistent Field Method and the Full Configuration Interaction Quantum Monte Carlo within a Super-CI Framework, with Application to Challenging Metal-Porphyrins. *Journal of Chemical Theory and Computation* **2016**, *12*, 1245–1258, PMID: 26808894.

(37) Li Manni, G.; Alavi, A. Understanding the Mechanism Stabilizing Intermediate Spin States in Fe(II)-Porphyrin. *J. Phys. Chem. A* **2018**, *122*, 4935–4947.

(38) Li Manni, G.; Kats, D.; Tew, D. P.; Alavi, A. Role of Valence and Semicore Electron Correlation on Spin Gaps in Fe(II)-Porphyrins. *Journal of Chemical Theory and Computation* **2019**, *15*, 1492–1497.

(39) Shavitt, I. Graph theoretical concepts for the unitary group approach to the many-electron correlation problem. *Int. J. Quantum Chem.* **1977**, *12*, 131.

(40) Shavitt, I. Matrix Element Evaluation in the Unitary Group Approach to the Electron Correlation Problem. *Int. J. Quantum Chem.* **1978**, *14 S12*, 5–32.

(41) Paldus, J. Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems. *J. Chem. Phys.* **1974**, *61*, 5321.

(42) Paldus, J. A pattern calculus for the unitary group approach to the electronic correlation problem. *Int. J. Quantum Chem.* **1975**, *9*, 165.

(43) Paldus, J. Unitary-group approach to the many-electron correlation problem: Relation of Gelfand and Weyl tableau formulations. *Phys. Rev. A* **1976**, *14*, 1620.

(44) Matsen, F. A. The unitary group formulation of the N-particle problem. *Int. J. Quantum Chem.* **1974**, *8*, 379.

(45) Louck, J. D. Recent progress toward a theory of tensor operators in the unitary groups. *Am. J. Phys.* **1970**, *38*, 3.

(46) Downward, M. J.; Robb, M. A. The computation of the representation matrices of the generators of the unitary group. *Theor. Chem. Acc.* **1977**, *46*, 129.

(47) Brooks, B. R.; Schaefer, H. F. The graphical unitary group approach to the electron correlation problem. Methods and preliminary applications. *J. Chem. Phys.* **1979**, *70*, 5092.

(48) Ma, D.; Li Manni, G.; Gagliardi, L. The generalized active space concept in multiconfigurational self-consistent field methods. *The Journal of Chemical Physics* **2011**, *135*, 044128.

(49) Ma, D.; Li Manni, G.; Olsen, J.; Gagliardi, L. Second-Order Perturbation Theory for Generalized Active Space Self-Consistent-Field Wave Functions. *J. Chem. Theory Comput.* **2016**, *12*, 3208–3213.

(50) Dobrautz, W.; Smart, S. D.; Alavi, A. Efficient formulation of full configuration interaction quantum Monte Carlo in a spin eigenbasis via the graphical unitary group approach. *The Journal of Chemical Physics* **2019**, *151*, 094104.

(51) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory*; John Wiley & Sons, Ltd: Chichester, England, 2000.

(52) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. Determinant Based Configuration Interaction Algorithms for Complete and Restricted Configuration Interaction Spaces. *J. Chem. Phys.* **1988**, *89*, 2185–2192.

(53) Grabenstetter, J. E.; Tseng, T. J.; Grein, F. Generation of genealogical spin eigenfunctions. *International Journal of Quantum Chemistry* **1976**, *10*, 143–149.

(54) Aquilante, F. et al. Molcas 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations Across the Periodic Table. *J. Comput. Chem.* **2016**, *37*, 506–541.

(55) Olsen, J. LUCIA, A correlation program written by J. Olsen with contributions from A. Koehn.

(56) Aidas, K. et al. The Dalton quantum chemistry program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 269–284.

(57) Olsen, J. A direct method to transform between expansions in the configuration state function and Slater determinant bases. *J. Chem. Phys.* **2014**, *141*, 034112.

(58) Fleig, T.; Olsen, J.; Marian, C. M. *J. Chem. Phys.* **2001**, *114*, 4775.

(59) Gel'fand, I. M.; Cetlin, M. L. Finite-dimensional representations of the group of orthogonal matrices. *Dokl. Akad. Nauk* **1950**, *71*, 1017, Amer. Math. Soc. Transl. 64, 116 (1967).

(60) Weyl, H. *The Classical Groups, Their Invariants and Representations*; Princeton U. P, Princeton, 1946.

(61) Shavitt, I. In *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*; Hinze, J., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1981; pp 51–99.

(62) Phung, Q. (Private Communication), 2018.