

Spin adaptation with determinant-based selected configuration interaction

Thomas Applencourt, Kevin Gasperich, Anthony Scemama

► **To cite this version:**

Thomas Applencourt, Kevin Gasperich, Anthony Scemama. Spin adaptation with determinant-based selected configuration interaction. [Research Report] CNRS; Université Paul Sabatier - Toulouse. 2018. hal-02468242

HAL Id: hal-02468242

<https://hal.archives-ouvertes.fr/hal-02468242>

Submitted on 5 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spin adaptation with determinant-based selected configuration interaction

Thomas Applencourt,¹ Kevin Gasperich,^{2,3} and Anthony Scemama^{4, a)}

¹⁾Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, Illinois 60439 USA

²⁾Computational Science Division, Argonne National Laboratory, Argonne, Illinois 60439 USA

³⁾Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260 USA

⁴⁾Laboratoire de Chimie et Physique Quantiques, Université de Toulouse, CNRS, UPS, France

Selected configuration interaction (sCI) methods, when complemented with a second order perturbative correction, provide near full configuration interaction (FCI) quality energies with only a small fraction of the Slater determinants of the FCI space. The selection of the determinants is often implemented in a determinant-based formalism, and therefore does not provide spin adapted wave functions. In other words, sCI wave functions are not eigenfunctions of the \hat{S}^2 operator. In some situations, having a spin adapted wave function is essential for the proper convergence of the method. We propose an efficient algorithm which, given an arbitrary determinant space, generates all the missing Slater determinants allowing one to obtain spin adapted wave functions while avoiding working with configuration state functions. For example, generating all the possible determinants with 6 up-spin and 6 down-spin electrons in 12 open shells takes 21 CPU cycles per generated Slater determinant. We also propose a modification of the denominators in the Epstein-Nesbet perturbation theory reducing significantly the non-invariance of the second order correction with respect to different values of the spin quantum number m_s . The computational cost of this correction is also negligible.

Keywords: Selected Configuration Interaction ; Spin adaptation ; Epstein-Nesbet perturbation theory

I. INTRODUCTION

In recent years, selected configuration interaction (sCI) methods have seen a resurgence in popularity,^{1–10} especially for the accurate calculation of electronic excitation energies.^{11–18} Balanced descriptions of excited states and dissociation curves require the wave functions to be spin adapted, i.e. eigenfunctions of the \hat{S}^2 operator. A natural option would be to reformulate sCI in terms of configuration state functions (CSF), but many codes were written in a determinant-based formulation, and opting for the CSF formalism would require a major re-writing of the software. Moreover, such a modification might increase the computational cost.^{19,20}

In the context of heat-bath selection, Holmes *et al* have improved the spin purity of the wave functions by introducing “time-reversal symmetry”,¹³ which consists of exchanging the spin labels of the electrons. However, when the number of open shells is large, time-reversal symmetry is not sufficient to generate all the required spin permutations among the open shells which would generate all the determinants of the corresponding CSFs.

A *space occupation pattern* (SOP) is a vector of occupation numbers of molecular orbitals. For example, the SOP (2, 1, 1, 1, 1) is made of two CSFs with coefficients μ and ν in the wave function, and 6 different Slater de-

terminants:

$$\begin{aligned} \begin{pmatrix} + \\ + \\ + \\ + \\ + \\ + \end{pmatrix} &= \mu \times \frac{1}{2} \left[\begin{pmatrix} \uparrow\downarrow \\ \downarrow\uparrow \\ \downarrow\uparrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} + \begin{pmatrix} \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} - \begin{pmatrix} \uparrow\downarrow \\ \downarrow\uparrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} - \begin{pmatrix} \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} \right] \\ &+ \nu \times \frac{\sqrt{3}}{6} \left[-2 \begin{pmatrix} \uparrow\downarrow \\ \downarrow\uparrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} + \begin{pmatrix} \uparrow\downarrow \\ \downarrow\uparrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} + \begin{pmatrix} \uparrow\downarrow \\ \downarrow\uparrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} \right. \\ &\quad \left. -2 \begin{pmatrix} \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} + \begin{pmatrix} \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} + \begin{pmatrix} \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \\ \uparrow\downarrow \end{pmatrix} \right]. \end{aligned} \quad (1)$$

A few years ago, Bytautas and Ruedenberg proposed a simple scheme to truncate large spin adapted wave functions while keeping the spin adaptation.²¹ By definition, all the determinants belonging to the same CSF have the same SOP, so the coefficients of the determinants with the same SOP are summed together to produce the so-called *space-product weights*, which are used to truncate the wave function. As spin coupling coefficients are implicitly included in the CI expansion, the truncated wave function is also an eigenfunction of \hat{S}^2 .

Following this idea, imposing spin adaptation in sCI methods can be done by

1. Identifying all the space occupation patterns of the determinants composing the variational space.
2. Generating all the determinants with imposed numbers of up-spin (\uparrow) and down-spin (\downarrow) electrons corresponding to these space occupation patterns.
3. Diagonalizing the Hamiltonian in this expanded determinant space.

^{a)}Electronic mail: scemama@irsamc.ups-tlse.fr

An efficient algorithm to achieve this procedure is presented in this paper, and then a modification to the Epstein-Nesbet perturbation expression is proposed. This modification introduces no additional cost, and it reduces the bias due to the lack of invariance with respect to the spin quantum number m_s . All the presented algorithms were implemented in the open-source *Quantum Package* software.²²

II. ALGORITHM

The wave function is expressed as

$$|\Psi\rangle = \sum_I c_I |D_I\rangle \quad (2)$$

Each Slater determinant D_I is represented as a Waller-Hartree double determinant,²³

$$D_I = d_i^\uparrow d_j^\downarrow, \quad (3)$$

the product of a determinant of \uparrow spin-orbitals with a determinant of \downarrow spin-orbitals. Such a representation can be encoded as a pair of bit strings $(\mathbf{d}_i, \mathbf{d}_j)$, where each bitstring is of length N_{orb} , the number of molecular orbitals. The spin-orbitals originate from a restricted Hartree-Fock or a complete active space (CAS) SCF calculation, such that the spatial parts of the spin-orbitals are common to the \uparrow and \downarrow spin-orbitals. Within a bit string, each bit corresponds to a spin-orbital; the bit is set to 1 if the orbital is occupied, and it is set to 0 if the orbital is empty. In low-level languages such as Fortran or C, a bit string may be stored as an array of N_{int} 64-bit integers, where

$$N_{\text{int}} = \left\lfloor \frac{N_{\text{orb}} - 1}{64} \right\rfloor + 1, \quad (4)$$

This representation allows for efficient determinant comparisons using bit-wise operation capabilities of modern processors²⁴ and will be convenient in the following.

All the CPU cycle measurements were performed on an Intel(R) Xeon(R) Gold 6140 CPU@ 2.30GHz with the GNU Fortran compiler 7.3.0, by reading the time stamp counter of the CPU with the `rdtsc` instruction.

A. Identification of the space occupation patterns

The SOP \mathbf{p}_I of determinant D_I , defined in Eq. (3), is a vector of integers defined as

$$[\mathbf{p}_I]_k = \begin{cases} 0 & \text{when the } k\text{-th orbital is unoccupied} \\ 1 & \text{when the } k\text{-th orbital is singly occupied} \\ 2 & \text{when the } k\text{-th orbital is doubly occupied} \end{cases} \quad (5)$$

```

function COMPUTE_PERMUTATIONS( $n, m$ )
  /*  $n$ : input, number of bits set to 1 */
  /*  $m$ : input, number of bits set to 0 */
  /*  $v$ : output, an array of permutations */
  /*  $u, t, t', t''$  and  $v$  are encoded in at least  $n+m+1$  bits */
   $k \leftarrow 0$ 
   $u \leftarrow (1 \ll n) - 1$ 
  while  $u < (1 \ll (n+m))$  do
     $v[k] \leftarrow u$ 
     $k \leftarrow k + 1$ 
     $t \leftarrow u \vee (u - 1)$ 
     $t' \leftarrow t + 1$ 
     $t'' \leftarrow ((\neg t \wedge t') - 1) \gg (\text{ctz}(u) + 1)$ 
     $u \leftarrow t' \vee t''$ 
  end while
  return  $v$ 
end function

```

FIG. 1. Anderson’s algorithm to generate all the patterns of n bits set to 1 in an integer of $n+m$ bits in lexicographic order. `ctz(i)` counts the number of trailing zeros, $i \ll n$ shifts i by n bits to the left, $i \gg n$ shifts i by n bits to the right, \wedge is the bit-wise **and** operator, and \vee is the bit-wise **or** operator.

If \mathbf{p}_I is encoded as a pair of bit strings $(\mathbf{p}_I^{(1)}, \mathbf{p}_I^{(2)})$, where $\mathbf{p}_I^{(1)}$ and $\mathbf{p}_I^{(2)}$ encode respectively the singly and doubly occupied orbitals, the SOP can be computed as

$$\begin{cases} \mathbf{p}_I^{(1)} & = \mathbf{d}_i \oplus \mathbf{d}_j \\ \mathbf{p}_I^{(2)} & = \mathbf{d}_i \wedge \mathbf{d}_j \end{cases} \quad (6)$$

where \oplus denotes the **xor** operator and \wedge denotes the **and** operator (see Fig. 2 for a pictorial representation).

Transforming all the determinants into a list of unique SOPs can be done in linear time if a hash value is associated with each SOP.²⁵ Hence, the time for this transformation is negligible.

B. Generating all the determinants

Given a SOP, one needs to generate all the possible excitations that can occur in the singly occupied molecular orbitals, keeping the numbers of \uparrow and \downarrow electrons fixed. One can remark that all the generated determinants will only differ by the singly occupied orbitals, so from now on we will consider a more compact representation: a bit string of $n_\uparrow + n_\downarrow$ bits, where n_\uparrow and n_\downarrow denote the numbers of \uparrow and \downarrow unpaired electrons. The bit is set to 1 when the orbital is occupied by an \uparrow electron, and 0 when it is occupied by a \downarrow electron. The indices of the singly occupied orbitals are kept in a look-up table \mathbf{m} for later use.

To generate all the determinants keeping the numbers of \uparrow and \downarrow electrons constant, we need to build all the possible bit strings with n_\uparrow bits set to 1 and n_\downarrow bits set to 0. This compact representation allows us to use Anderson’s algorithm,²⁶ which generates all the patterns of n_\uparrow bits set to 1 in a bit string of length $n_\uparrow + n_\downarrow$ in

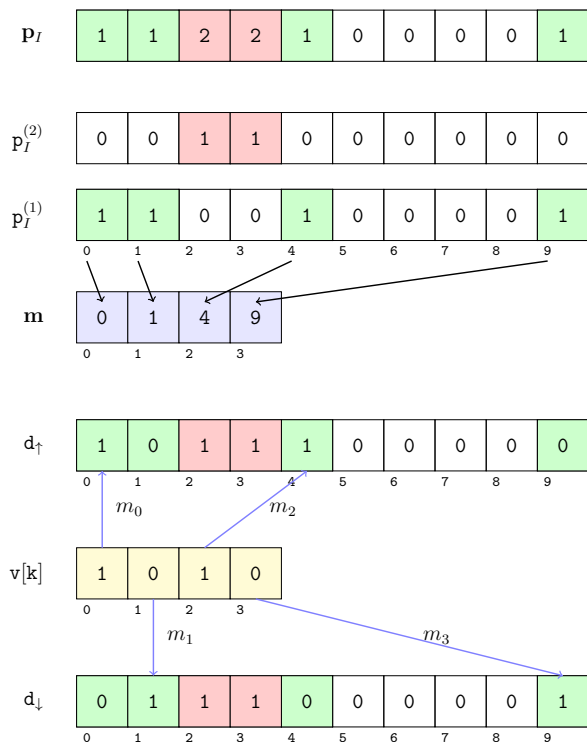


FIG. 2. The SOP \mathbf{p}_I is encoded as in Eq. (6). Singly and doubly occupied orbitals are represented respectively in green and red. The list of indices \mathbf{m} of the singly occupied orbitals is built (in blue), and this mapping is re-used to build the determinants from permutations (yellow) generated by Anderson’s algorithm.

lexicographical order. For example, with $n_{\uparrow} = 2$ and $n_{\downarrow} = 2$ it produces the sequence (0011, 0101, 0110, 1001, 1010, 1100).

The algorithm proceeds as follows. The integer u is initialized with $2^{n_{\uparrow}+1} - 1$, namely the smallest possible unsigned integer with n_{\uparrow} bits set to 1. Then, the following steps are iterated until u becomes greater than $2^{n_{\uparrow}+n_{\downarrow}} - 1$:

1. Set all the least significant 0 bits of u to 1, add 1 and store the result in \mathbf{t}' . The least significant 1 of \mathbf{t}' marks the position in u of the most significant 0 that should be changed into a 1.
2. The position of the least significant 1 of u is identified by counting the number of trailing zeros in u . This 1 should be changed into a 0.
3. At the right of this position, the least significant 0’s should be changed to 1’s such that the total number of 1’s is the same as in u .

The corresponding pseudo-code is presented in Fig. 1. On average, one loop cycle executes in 8.2 CPU cycles.

Fig. 2 gives a pictorial description of the data structures used to generate a determinant. To build a generated determinant ($d_{\uparrow}, d_{\downarrow}$) from a permutation u , one needs to

1. Fill the doubly occupied orbitals by setting both d_{\uparrow} and d_{\downarrow} equal to $p_I^{(2)}$.
2. Iterate over the bits of u . If the k -th bit is set to 1, set the m_k -th orbital of d_{\uparrow} to 1, otherwise set the m_k -th orbital of d_{\downarrow} to 1.

C. Further optimizations

As a first optimization, instead of creating each determinant from the permutation as shown in Fig. 2, all the determinants can be generated iteratively by considering only the orbitals that have changed from the previously generated determinant. This avoids always setting all the $n_{\uparrow} + n_{\downarrow}$ bits in the bit strings. The integer obtained by $v[k-1] \oplus v[k]$ has bits set to 1 at the positions where the bits differ between $v[k-1]$ and $v[k]$. The positions of these bits can be found in a few cycles by

1. Counting the number of trailing zeros. This gives the position of the least significant 1.
2. Setting the least significant 1 to 0 using $v[k] \leftarrow v[k] \wedge (v[k] - 1)$.

and iterating until $v[k] = 0$.

A second optimization is to consider time-reversal symmetry. When $n_{\uparrow} = n_{\downarrow}$, one can remark that $v[\mathbf{n}_{\text{det}} - 1 - \mathbf{k}] = \neg v[\mathbf{k}]$, where \mathbf{n}_{det} is the number of determinants generated:

$$\mathbf{n}_{\text{det}} = \frac{(n_{\uparrow} + n_{\downarrow})!}{n_{\uparrow}!n_{\downarrow}!} \quad (7)$$

Hence, it suffices to iterate over the first half of the permutations of Anderson’s algorithm, and generate two determinants per iteration.

III. SHIFTED EPSTEIN-NESBET DENOMINATORS

Let us consider a real-valued normalized spin-adapted wave function with energy E , expressed as

$$|\Psi\rangle = \sum_{i \in \mathcal{I}} c_I |D_I\rangle \quad (8)$$

which is an eigenfunction of the Hamiltonian projected in the internal space of determinants \mathcal{I} . The variance of the energy associated with this function is

$$\sigma^2 = \langle \Psi | \hat{H}^2 | \Psi \rangle - \langle \Psi | \hat{H} | \Psi \rangle^2. \quad (9)$$

Inserting the resolution of the identity for \hat{H}^2 gives an approximation of the variance truncated to the full configuration interaction (FCI) space

$$\sigma^2 = \sum_{\alpha \in \mathcal{F}} \langle \Psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \Psi \rangle - E^2 \quad (10)$$

where \mathcal{F} denotes a complete set of arbitrary orthonormal basis functions, $|\alpha\rangle$, spanning the FCI space. The FCI space can be split in three subspaces:

- The internal space \mathcal{I} .
- The external space \mathcal{E} which is the subset of functions, $|\alpha\rangle$ which don't belong to \mathcal{I} , and for which, $\langle\alpha|\hat{H}|\Psi\rangle \neq 0$.
- The rest of the FCI space.

\hat{H} is symmetric and $|\Psi\rangle$ is real, so Eq. (10) can be rewritten as

$$\sigma^2 = \sum_{D_I \in \mathcal{I}} \langle D_I | \hat{H} | \Psi \rangle^2 + \sum_{\alpha \in \mathcal{E}} \langle \alpha | \hat{H} | \Psi \rangle^2 - E^2. \quad (11)$$

As $|\Psi\rangle$ is an eigenfunction of \hat{H} projected in \mathcal{I} ,

$$\langle D_I | \hat{H} | \Psi \rangle^2 = (E \langle D_I | \Psi \rangle)^2 = E^2 c_I^2, \quad (12)$$

and as $|\Psi\rangle$ is normalized, one obtains

$$\sigma^2 = \sum_{\alpha \in \mathcal{E}} \langle \alpha | \hat{H} | \Psi \rangle^2. \quad (13)$$

The variance of the energy does not depend on the particular choice of the functions $|\alpha\rangle$, as long as they constitute an orthonormal set of functions spanning the space \mathcal{E} . Moreover, the variance of the energy is equal for degenerate wave functions with different spin quantum numbers m_s . Hence, one can choose equivalently the $|\alpha\rangle$'s to be Slater determinants or CSFs.

The Epstein-Nesbet (EN) second-order perturbative contribution to the energy is given by

$$E_{\text{PT2}} = \sum_{\alpha \in \mathcal{E}} \frac{\langle \alpha | \hat{H} | \Psi \rangle^2}{E - \langle \alpha | \hat{H} | \alpha \rangle}. \quad (14)$$

This equation can be seen as a weighted sum of the different terms involved in the expression of the variance. However, the weights differ depending on the choice of $|\alpha\rangle$. Also, when a basis of Slater determinants is chosen, this expression is not invariant with respect to the choice of m_s , and this is not desirable.

A way to cure the lack of invariance with respect to m_s is to impose all the weights to be the same for all the determinants belonging to the same CSF. But as the same determinant can appear in the expression of multiple CSFs, Davidson proposed to use a modified EN zeroth-order Hamiltonian formed from diagonal elements averaged over Slater determinants belonging to the same SOP. This idea was implemented a long time ago in the MELD program,^{27,28} and also in SCIEL.²⁹

This modified zeroth order Hamiltonian implies that the weight is the same for all the terms associated with determinants belonging to the same SOP. This can be

done by inserting a determinant-specific energy shift ϵ_α to the diagonal element at the denominator

$$E_{\text{PT2}} = \sum_{\alpha \in \mathcal{E}} \frac{\langle \alpha | \hat{H} | \Psi \rangle^2}{E - (\langle \alpha | \hat{H} | \alpha \rangle + \epsilon_\alpha)}. \quad (15)$$

with

$$\epsilon_\alpha = E_\alpha - \langle \alpha | \hat{H} | \alpha \rangle \quad (16)$$

where the shift is chosen to be

$$E_\alpha = \min_{\beta \in \text{SOP}(\alpha)} \langle \beta | \hat{H} | \beta \rangle. \quad (17)$$

where the $|\beta\rangle$ determinants run over all the determinants belonging to the same SOP as $|\alpha\rangle$. This choice of E_α is not the same as Davidson's, but it also gives the same weight for all the values of m_s . Although the generation of all the determinants is extremely fast, using this approximation can become expensive since it requires the computation of all diagonal elements $\langle \beta | \hat{H} | \beta \rangle$ for each $|\alpha\rangle$.

To circumvent this problem, one can remark that for the majority of the contributions to

$$\langle \alpha | \hat{H} | \Psi \rangle = \sum_I c_I \langle \alpha | \hat{H} | D_I \rangle, \quad (18)$$

$|\alpha\rangle$ is doubly excited with respect to $|D_I\rangle$. We now consider that $|D_I\rangle$ is the determinant with the lowest energy among all the determinants sharing the same SOP. For all the other determinants $|D_J\rangle$ belonging to the same SOP as $|D_I\rangle$ and doubly excited with respect to $|D_I\rangle$, one can define a double excitation operator

$$\hat{T}_{I \rightarrow J} | D_I \rangle = | D_J \rangle. \quad (19)$$

Remarking that

$$\langle \hat{T}_{I \rightarrow J} D_I | \hat{H} | \hat{T}_{I \rightarrow J} \alpha \rangle \begin{cases} \pm \langle \alpha | \hat{H} | D_I \rangle & \text{if } \hat{T}_{I \rightarrow J} |\alpha\rangle \neq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (20)$$

the contributions connected by \hat{H} to $|D_J\rangle$ are the $\hat{T}_{I \rightarrow J} |\alpha\rangle$ which have a diagonal element which will be shifted by

$$E_\alpha = \langle \hat{T}_{I \rightarrow J} \alpha | \hat{H} | \hat{T}_{I \rightarrow J} \alpha \rangle - \langle \alpha | \hat{H} | \alpha \rangle \quad (21)$$

This quantity may be approximated by

$$E_J = \langle D_J | \hat{H} | D_J \rangle - \langle D_I | \hat{H} | D_I \rangle, \quad (22)$$

and the approximate shifts E_J can be precomputed. If $|\alpha\rangle$ is connected to multiple $|D_I\rangle$'s, we take the energy shift associated with the $|D_I\rangle$ with largest associated $|c_I|$. As our implementation generates the $|\alpha\rangle$'s with no duplicates from the $|D_I\rangle$ sorted by decreasing $|c_I|$,³⁰ the use of the shift can be made at no cost.

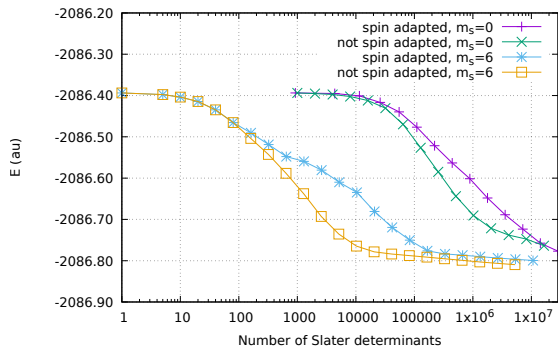


FIG. 3. Variational energy of dissociated Cr_2 as a function of the number of selected Slater determinants in the wave function expansion.

IV. NUMERICAL TESTS

A. Open-shell toy problem

To test our implementation with a large number of open shells, we have prepared model wave functions for the dissociated Chromium dimer in its 13-et state, separated by a distance of 100 Å, using the def2-SVP basis.³¹ At such a large distance, each Chromium atom is in its high-spin state with 6 unpaired electrons. Two equivalent wave functions are built to initialize the sCI calculation:

- The $m_s = 6$ wave function, which is a single determinant with 30 \uparrow and 18 \downarrow electrons.
- The 13-et $m_s = 0$ wave function with 24 \uparrow and 24 \downarrow electrons, which contains 924 determinants.

The system is composed of 62 molecular orbitals, so for this simple case $N_{\text{int}} = 1$. The orbitals were obtained at the restricted open-shell Hartree-Fock (ROHF) level for $m_s = 6$. The selection is performed in the valence FCI space, with 20 frozen electrons.

The $m_s = 0$ wave function was initialized by taking the same SOP as the one of the single determinant of the $m_s = 6$ wave function, and generating all the possible determinants using the algorithm presented in this paper. The generation of the 924 determinants was done in $\sim 19\,200$ CPU cycles (~ 8 microseconds), i.e. 21 cycles per generated determinant. The Hamiltonian was diagonalized and we checked that the lowest state with $\langle \hat{S}^2 \rangle = 42$ had the exact same energy as the $m_s = 6$ single determinant.

For the $m_s = 6$ wave function, we have run a CIPSI selection constraining or not the wave function to be spin-adapted. As expected, for the same energy the number of determinants is increased when spin adaptation is imposed. Then, we have run the CIPSI selection for the $m_s = 0$ wave function imposing the spin adaptation. The convergence of the energy is plotted in Fig. 3 and 4 as a function of the number of determinants and of the number of SOP. From Fig. 4, it is striking that the $m_s = 0$

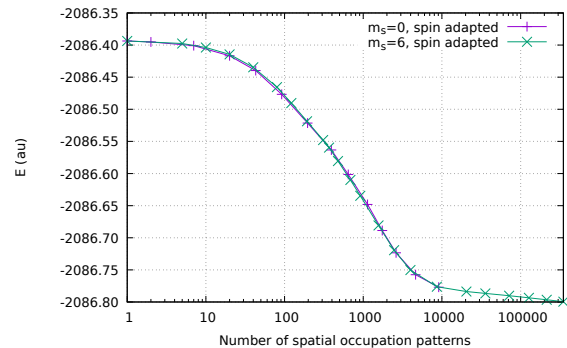


FIG. 4. Variational energy of dissociated Cr_2 as a function of the number of selected SOP.

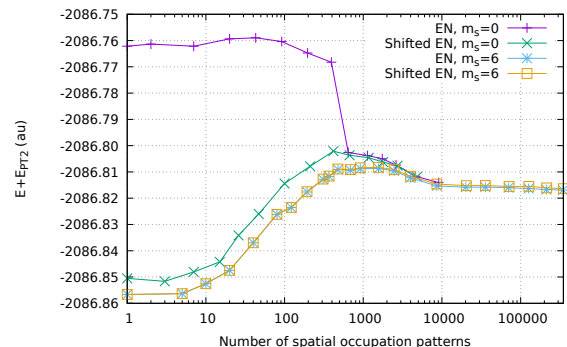


FIG. 5. Variational energy with second order perturbative correction of dissociated Cr_2 as a function of the number of selected SOP, using the EN denominators or the shifted EN denominators.

and $m_s = 6$ wave functions are indeed equivalent. This example exhibits the fact that having a large number of determinants with small weights is not always characteristic of a multi-reference character. Moreover, the number of determinants is not a relevant criterion for the quality of a wave function, as opposed to the number of CSFs. However, using SOPs appears as a cheap alternative to CSFs when working in the determinant framework.

Fig. 5 shows that the EN $E_{\text{PT}2}$ values are very different between $m_s = 6$ and $m_s = 0$ when the number of SOP is less than 1000. The shifted EN and the EN values give almost identical energy curves for $m_s = 6$; however, for $m_s = 0$ the shifted EN fixes the incorrect behavior of the curve for small numbers of SOPs. For larger numbers of SOPs, the two $m_s = 0$ curves join before converging to the $m_s = 6$ curve. The joining of the shifted EN and EN curves signifies that all the determinants of the external space with low energies have been included in the internal space. In addition, the E term in the denominator becomes more negative as plotted in Fig. 4, so all the denominators tend to be large enough that the use of the energetic shift becomes less and less important.

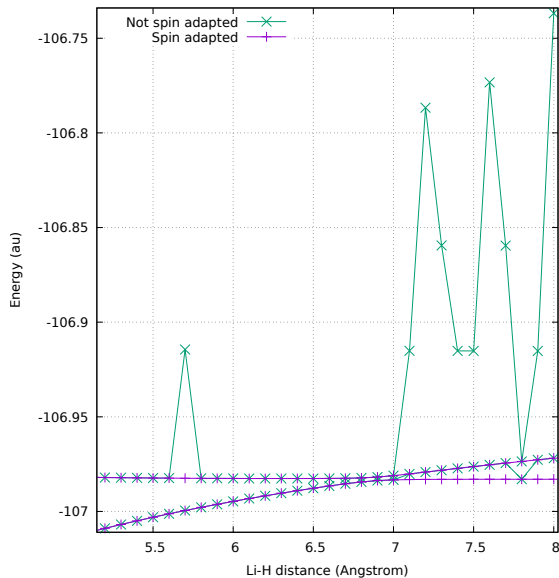


FIG. 6. Avoided crossing of LiF, with and without imposing spin adaptation.

B. Avoided crossing of LiF

The avoided crossing between the ionic and neutral ${}^1\Sigma^+$ states of LiF is a common benchmark for correlated methods, as the location of the crossing is highly sensitive to the amount of correlation. At large distances, the lowest triplet state is very close in energy to the singlet states. If the wave function is not spin adapted, the triplet state will mix with the singlets during the selection, and the convergence of the CIPSI calculations to the correct states is not guaranteed.

To show the importance of spin adaptation for such a problem, we have reported in Fig. 6 the potential energy curve of the two lowest singlet states of LiF computed with and without imposing spin adaptation. For all the distances, the CIPSI calculations were run blindly (with no user interaction), starting with the CAS-SCF(2,2)/aug-cc-pVDZ wave functions of both states (four determinants). Only the lowest molecular orbital was frozen, corresponding to the $1s$ orbital of the Fluorine atom. The calculations were stopped when the second-order perturbative contribution was below $0.1 mE_h$ or when the number of determinants reached 4 million.

Fig. 6 shows that for large distances, without spin adaptation there are multiple erratic points for which the two obtained states are not the desired ones. This curve also shows that all the points obtained with spin adaptation converged to the correct states, giving a smooth potential energy curve.

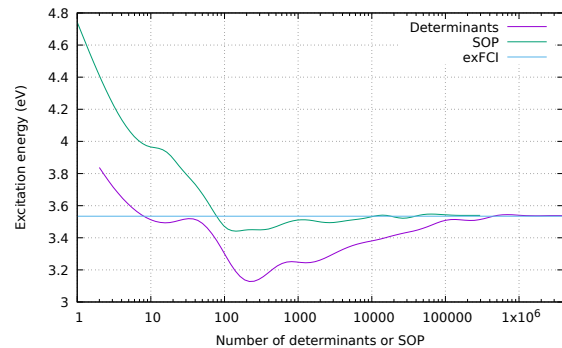


FIG. 7. Adiabatic transition energy of formaldehyde, computed with fixed numbers of determinants and fixed numbers of SOP.

C. Adiabatic transition energy of formaldehyde

The last example we present is the calculation of the adiabatic transition energy of the formaldehyde molecule with the aug-cc-pVDZ basis set. The ground state is well described by a single determinant, and the excited state is an open shell singlet obtained by a ${}^1(n \rightarrow \pi^*)$ single excitation, so its main CSF contains two determinants. The geometries of the ground and excited states were taken from Loos *et al.*¹⁴. For both geometries, a preliminary CIPSI calculation was run to produce state-averaged natural orbitals, in order to work with molecular orbitals of comparable quality for both states. Hence, the rate of convergence of the energy with respect to the number of selected CSFs is expected to be comparable for both states. However, we expect the rate of convergence of the energy with respect to the number of selected determinants to be different.

For each state a state-specific CIPSI calculation was run at its equilibrium geometry. For the excited state, the run was initiated using a wave function with the two determinants of the reference CSF, and a state-following approach based on the maximum overlap with the initial guess was used during the Davidson diagonalizations to avoid collapsing to the ground state.

The qualitative difference between the two states (single determinant *vs* two-determinant open-shell singlet) makes the computation of the energy difference inaccurate if the energy difference is calculated for the same number of selected determinants. However, one can compute the energy difference for the same number of selected SOP, which is expected to be consistent with the use of the same number of CSFs.

Fig. 7 plots the adiabatic transition energy computed taking the energies of both states with the same number of selected determinants, and taking the energies of both states with the same number of SOP. As the two runs were independent, the number of selected determinants and SOPs were different for the two states so a cubic spline interpolation was used to compute the energies at arbitrary numbers of determinants. This fig-

ure shows that the adiabatic transition energy converges much faster to 3.53 eV using the SOP criterion, a value close to the experimental value of 3.50 eV.^{32,33}

V. CONCLUSION

We have presented a general algorithm to complement an arbitrary wave function with all the required Slater determinants to obtain eigenstates of the \hat{S}^2 operator when the Hamiltonian is diagonalized, with a negligible computational overhead. This spin adaptation is introduced after the selection of determinants in the selected CI algorithm. The presented examples have illustrated different situations where spin adaptation is important within sCI. When comparing wave functions, considering the number of CSFs is more relevant than the number of Slater determinants and considering SOPs allows one to stay in the determinant framework while still benefiting from the consistency brought by CSFs. Finally, we would like to emphasize that this spin adaptation procedure can be applied to any selected CI method: CIPSI, heat-bath CI, machine learning CI, Monte Carlo CI, *etc.*

ACKNOWLEDGMENTS

The authors gratefully acknowledge Sean Eron Anderson for creating the *Bit Twiddling Hacks* web page. This work was performed using HPC resources from CALMIP (Toulouse) under allocation 2018-0510 and from GENCI-TGCC (Grant 2018-A0040801738). KG acknowledges support from grant number CHE1762337 from the U.S. National Science Foundation. This research used resources of the Argonne Leadership Computing Facility, which is a U.S. Department of Energy Office of Science User Facility operated under contract DE-AC02-06CH11357.

Fortran/C/++ implementations of the Fig. 2 algorithm are available at <https://github.com/Tapplencourt/occ2det>.

¹J. Greer, *Journal of Computational Physics* **146**, 181 (1998).

²P. Stampfuß and W. Wenzel, *The Journal of Chemical Physics* **122**, 024110 (2005).

³L. Bytautas and K. Ruedenberg, *Chemical Physics* **356**, 64 (2009).

⁴G. H. Booth, A. J. W. Thom, and A. Alavi, *The Journal of Chemical Physics* **131**, 054106 (2009).

⁵E. Giner, A. Scemama, and M. Caffarel, *Canadian Journal of Chemistry* **91**, 879 (2013).

⁶R. J. Buenker, R. A. Phillips, S. Krebs, H.-P. Liebermann, A. B. Alekseyev, and P. Funke, *Theoretical Chemistry Accounts* **133** (2014), 10.1007/s00214-014-1468-7.

⁷A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *Journal of Chemical Theory and Computation* **12**, 3674 (2016).

⁸Y. Ohtsuka and J. ya Hasegawa, *The Journal of Chemical Physics* **147**, 034102 (2017).

⁹J. P. Coe, *Journal of Chemical Theory and Computation* **14**, 5739 (2018).

¹⁰A. Scemama, Y. Garniron, M. Caffarel, and P.-F. Loos, *Journal of Chemical Theory and Computation* **14**, 1395 (2018).

¹¹J. P. Coe and M. J. Paterson, *The Journal of Chemical Physics* **139**, 154103 (2013).

¹²J. B. Schriber and F. A. Evangelista, *Journal of Chemical Theory and Computation* **13**, 5354 (2017).

¹³A. A. Holmes, C. J. Umrigar, and S. Sharma, *The Journal of Chemical Physics* **147**, 164111 (2017).

¹⁴P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, *Journal of Chemical Theory and Computation* **14**, 4360 (2018).

¹⁵A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.-F. Loos, *The Journal of Chemical Physics* **149**, 034108 (2018).

¹⁶M. Dash, S. Moroni, A. Scemama, and C. Filippi, *Journal of Chemical Theory and Computation* **14**, 4176 (2018).

¹⁷A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma, and P. M. Zimmerman, *The Journal of Physical Chemistry A* **122**, 2714 (2018).

¹⁸P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, "Reference Energies for Double Excitations," (2018), [Online; accessed 13. Dec. 2018].

¹⁹P. Knowles and N. Handy, *Chemical Physics Letters* **111**, 315 (1984).

²⁰J. Olsen, B. O. Roos, P. Joergensen, and H. J. A. Jensen, *The Journal of Chemical Physics* **89**, 2185 (1988).

²¹L. Bytautas and K. Ruedenberg, in *ACS Symposium Series* (American Chemical Society, 2007) pp. 103–123.

²²"Quantum package," (2018).

²³R. Pauncz, *International Journal of Quantum Chemistry* **35**, 717 (1989).

²⁴A. Scemama and E. Giner, arXiv (2013), 1311.6244.

²⁵D. Bitton and D. J. DeWitt, *ACM Trans. Database Syst.* **8**, 255 (1983).

²⁶"Bit twiddling hacks," (2018), accessed Nov 15 2018.

²⁷E. Davidson, L. Nitzche, and L. McMurchie, *Chemical Physics Letters* **62**, 467 (1979).

²⁸P. M. Kozlowski and E. R. Davidson, *The Journal of Chemical Physics* **100**, 3672 (1994).

²⁹R. Caballol, J. P. Malrieu, J. P. Daudey, and O. Castell, "Scienc program," (1998).

³⁰Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, *The Journal of Chemical Physics* **147**, 034101 (2017).

³¹F. Weigend and R. Ahlrichs, *Physical Chemistry Chemical Physics* **7**, 3297 (2005).

³²D. J. Clouthier and D. A. Ramsay, *Annual Review of Physical Chemistry* **34**, 31 (1983).

³³C. Angeli, S. Borini, L. Ferrighi, and R. Cimbriglia, *The Journal of Chemical Physics* **122**, 114304 (2005).