Calculation matrix elements in XMS-CASPT2

Peter John Cherry
Shiozaki group meeting
September 6th 2017

Derivatives of XMS-CASPT2 wavefunctions

- The XMS-CASPT2 energy is not minimized with respect to the CI and orbital coefficients obtained in the CASSCF calculation.
- This makes differentiating it with respect to perturbations of the Hamiltonian difficult.
- Accordingly, a Lagrangian is defined, which is minimized with respect to these variables:

$$L(\mathbf{t}, \mathbf{c}, \mathbf{z}, \boldsymbol{\lambda}) = E(\mathbf{t}, \mathbf{c}) + \mathbf{z}^\dagger \mathbf{g}(\mathbf{c}) + \boldsymbol{\lambda}^\dagger \mathbf{g}'(\mathbf{t})$$

- Typically, differentiating this Lagrangian with respect to some perturbative parameter is much easier than differentiating the XMS-CASPT2 energy.

CI derivatives

$$\frac{\partial L}{\partial \tilde{c}_{Q,K}} = \sum_M \langle \tilde{M}|\hat{H}|K\rangle R^*_{MP} R_{QP} + \sum_N \langle K|\hat{H}|\tilde{N}\rangle R^*_{QP} R_{KP}$$

$$+ \sum_L R^*_{LP} \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{H}|K\rangle R_{QP} + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{H}|\tilde{N}\rangle R_{NP} \right)$$

$$- E_s \sum_L \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_L \left(\sum_M \langle \tilde{M}|\hat{\lambda}^\dagger_{LM}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{\lambda}^\dagger_{LQ}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_{rs} \Big(\langle K|\hat{E}_{rs}|\tilde{M}\rangle + \langle \tilde{M}|\hat{E}_{rs}|K\rangle \Big) [W_M \mathbf{g}(\mathbf{d}^{(2)}) - N_M \mathbf{f}]_{rs}$$

Calculation of matrix elements for many electron operators

- Need a generic approach for evaluating terms of the form

$$\sum_{\substack{ijkl\\wxyz}} \sum_{J} T^{\dagger}_{ijkl} g_{wxyz} q_{uv} \langle I | a_i a_j a^{\dagger}_k a^{\dagger}_l a^{\dagger}_w a^{\dagger}_x a_y a_z a^{\dagger}_u a_v | J \rangle c_J$$

- SMITH3 generates code of evaluating these integrals.

- Jae developed and implemented and algorithm such that the most costly term is

$$\sum_{ijklmn}^{act} \sum_{J} \langle I | a^{\dagger}_i a^{\dagger}_j a^{\dagger}_k a_l a_m a_n | J \rangle c_J A_{ijklmn}$$

- This is highly efficient, but unfortunately, it is not as well suited to the cases with spin-flipping interactions.

Derivatives of XMS-CASPT2 wavefunctions

- In a relativistic framework, magnetic properties are often obtained via the principal of minimal coupling:

$$\hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}} + i\mathbf{A}$$

- Here is describes the perturbation.

CI derivatives

$$\frac{\partial L}{\partial \tilde{c}_{Q,K}} = \sum_M \langle \tilde{M}|\hat{H}|K\rangle R^*_{MP} R_{QP} + \sum_N \langle K|\hat{H}|\tilde{N}\rangle R^*_{QP} R_{KP}$$

$$+ \sum_L R^*_{LP} \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{H}|K\rangle R_{QP} + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{H}|\tilde{N}\rangle R_{NP} \right)$$

$$- E_s \sum_L \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_L \left(\sum_M \langle \tilde{M}|\hat{\lambda}^\dagger_{LM}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{\lambda}^\dagger_{LQ}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_{rs} \Big(\langle K|\hat{E}_{rs}|\tilde{M}\rangle + \langle \tilde{M}|\hat{E}_{rs}|K\rangle \Big) [W_M \mathbf{g}(\mathbf{d}^{(2)}) - N_M \mathbf{f}]_{rs}$$

CI derivatives

$$\frac{\partial L}{\partial \tilde{c}_{Q,K}} = \sum_M \langle \tilde{M}|\hat{H}|K\rangle R^*_{MP} R_{QP} + \sum_N \langle K|\hat{H}|\tilde{N}\rangle R^*_{QP} R_{KP}$$

$$+ \sum_L R^*_{LP} \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{H}|K\rangle R_{QP} + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{H}|\tilde{N}\rangle R_{NP} \right)$$

$$- E_s \sum_L \left(\sum_M \langle \tilde{M}|\hat{T}^\dagger_{LM}\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{T}^\dagger_{LQ}\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_L \left(\sum_M \langle \tilde{M}|\hat{\lambda}^\dagger_{LM}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LQ}|K\rangle + \sum_N \langle K|\hat{\lambda}^\dagger_{LQ}(\hat{f} - \epsilon^{(0)}_L + \epsilon_s)\hat{T}_{LN}|\tilde{N}\rangle \right)$$

$$+ \sum_{rs} \Big(\langle K|\hat{E}_{rs}|\tilde{M}\rangle + \langle \tilde{M}|\hat{E}_{rs}|K\rangle \Big) [W_M \mathbf{g}(\mathbf{d}^{(2)}) - N_M \mathbf{f}]_{rs}$$

Calculation of matrix elements for many electron operators

- Main differences in new code:
  - Takes advantage of time reversal symmetry in the relativistic case.
  - Does not express the matrix element as a sum of normal ordered terms, instead uses alternating order, e.g.,

$$\langle M|a_i^\dagger a_j a_k^\dagger a_m a_n^\dagger a_o|N\rangle \qquad \underline{\textit{not}} \qquad \langle M|a_i^\dagger a_j^\dagger a_k^\dagger a_m a_n a_o|N\rangle$$

  - Operators may be repeatedly reordered so as to take advantage of symmetry conditions.
  - Currently does not generate code; instead uses a generic algorithm.

Method outline

Step 1 : Read in tensors corresponding to operators represented in the molecular orbital basis.

Step 2 : Use constraints and symmetry conditions to determine unique blocks.

Step 3 : Determine all possible contractions, $A_{ijk\ldots}$, of these tensors.

Step 4 : Calculate γ^{IJ} matrices for the relevant spin sector.

Step 5 : Contract γ^{IJ} matrices with the relevant $A_{ijk\ldots}$.

Guiding principal : Relativistic and non-relativistic cases should be treated using the same basic algorithm.

Methodological outline

Step 1a : Read molecular orbital representations of operators, their symmetry conditions and constraints.

Step 1b : Flatten pairs of indices, e.g,

$$T_{ij,kl} \rightarrow T'_{(ij),(kl)} = T'_{ab}$$

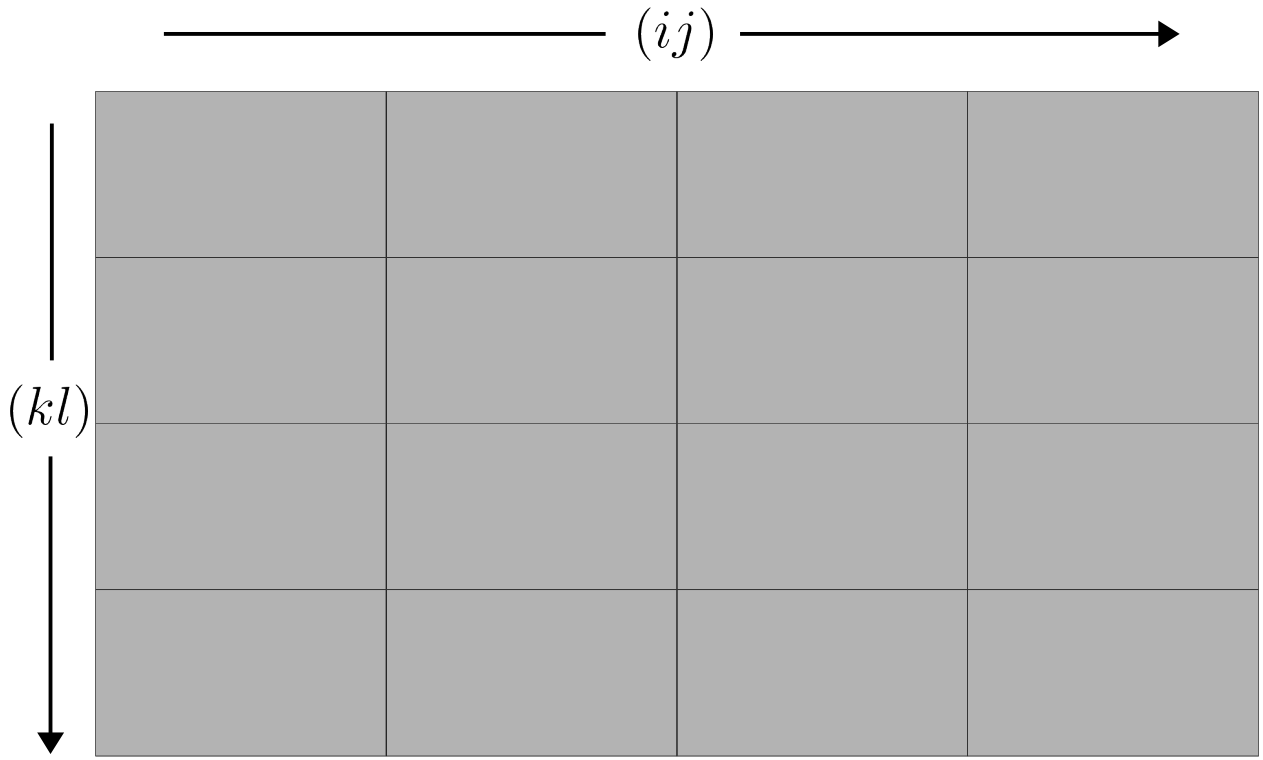
Step 1c : Decompose the matrix into blocks, and using the constraints and symmetry relations to determine which ones vanish or are equivalent, e.g.,

$$T_{(ij),(kl)} = T_{(kl),(ij)} = T_{(lk),(ji)} = T_{(ji),(kl)}$$

Step 1d : Construct a map which can be used to identify only those matrix elements which are necessary.

Application of block symmetry

Example : CASPT2 perturbation coefficients $T_{(ij),(kl)}$.



Possible transitions

$$closed \rightarrow active$$
$$closed \rightarrow virtual$$
$$active \rightarrow active$$
$$active \rightarrow virtual$$

Application of block symmetry

Example : CASPT2 perturbation coefficients $T_{(ij),(kl)}$.

	$c \to a$	$c \to v$	$a \to a$	$a \to v$
$c \to a$				
$c \to v$				
$a \to a$				
$a \to v$				

Possible transitions

$$closed \to active$$
$$closed \to virtual$$
$$active \to active$$
$$active \to virtual$$

Application of block symmetry

Example : CASPT2 perturbation coefficients $T_{(ij),(kl)}$.

	$c \rightarrow a$	$c \rightarrow v$	$a \rightarrow a$	$a \rightarrow v$
$c \rightarrow a$				
$c \rightarrow v$				
$a \rightarrow a$				
$a \rightarrow v$				

Possible transitions

$closed \rightarrow active$
$closed \rightarrow virtual$
$active \rightarrow active$
$active \rightarrow virtual$

Application of block symmetry

Example : CASPT2 perturbation coefficients $T_{(ij),(kl)}$.

	$c \rightarrow a$	$c \rightarrow v$	$a \rightarrow a$	$a \rightarrow v$
$c \rightarrow a$				
$c \rightarrow v$				
$a \rightarrow a$				
$a \rightarrow v$				

Possible transitions

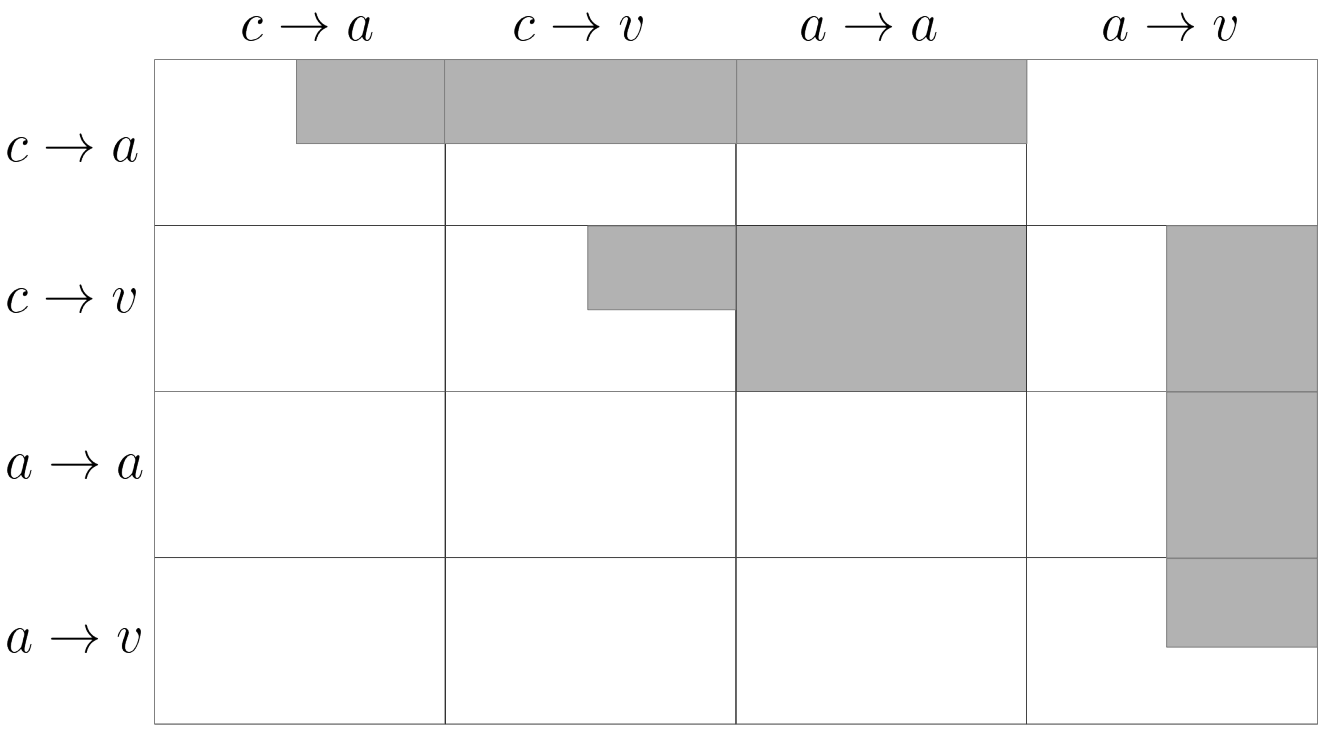
$closed \rightarrow active$
$closed \rightarrow virtual$
$active \rightarrow active$
$active \rightarrow virtual$

Block specific permutation symmetry

- Within each block the ranges of the indices are constrained to specific ranges.
- For a given block, it may be that two indexes cannot be interchanged, as their ranges are different.
- Hence the permutation symmetry varies between blocks.

Application of symmetry

Example : CASPT2 perturbation coefficients $T_{(ij),(kl)}$.



Possible transitions

$closed \to active$
$closed \to virtual$
$active \to active$
$active \to virtual$

Method outline

Step 1 : Read in tensors corresponding to operators represented in the molecular orbital basis.

Step 2 : Use constraints and symmetry conditions to determine unique blocks.

Step 3 : Determine all possible contractions, $A_{ijk\ldots}$, of these tensors.

Guiding principal : Relativistic and non-relativistic cases should be treated using the same basic algorithm.

→ *Relativistic wavefunctions just have twice as many blocks (from spin), and extra symmetry relations.*

Contraction of tensor blocks

Step 3a : Contract over all possible pairs of indexes, e.g.,

$$\hat{H}\hat{T} \rightarrow \sum_{wxyz} \sum_{ijkl} H_{wxyz} a_w^\dagger a_x^\dagger a_y a_z T_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$$

$$A^{HT,st}_{abcdef} = \sum_{wxyz} \sum_{ijkl} H_{wxyz} T_{ijkl} \delta_{st}$$

$$s := w, x, i, \text{ or } j$$
$$t := y, z, k \text{ or } l$$

Step 3b : Use the block range constraints to rule out possible contractions, e.g.

$$A^{HT,st}_{abcdef} = \begin{cases} A^{HT,st}_{abcdef} & \text{if} \quad rng(s) = rng(t) \\ 0 & \text{otherwise} \end{cases}$$

Method outline

Step 1 : Read in tensors corresponding to operators represented in the molecular orbital basis.

Step 2 : Use constraints and symmetry conditions to determine unique blocks.

Step 3 : Determine all possible contractions, $A_{ijk\ldots}$, of these tensors.

Step 4 : Calculate γ^{IJ} matrices for the relevant spin sector.

Guiding principal : Relativistic and non-relativistic cases should be treated using the same basic algorithm.

→ *Relativistic wavefunctions just have twice as many blocks (from spin), and extra symmetry relations.*

Calculation of $\gamma^I_{ijk\ldots}$

Step 4a : First take creation and annihilation operators in order specified by input operator, e.g.,

$$\langle I|a_l a_i^\dagger a_m a_j^\dagger a_n a_k^\dagger|J\rangle$$

Step 4b : Rewrite in normal order:

$$\begin{aligned} = \langle I|a_i^\dagger a_j^\dagger a_k^\dagger a_l a_m a_n|J\rangle + (s_{m'n'}\delta_{m'n'} + \ldots)\langle I|a_{i'}^\dagger a_{j'}^\dagger a_{k'} a_{l'}|J\rangle \\ +(s_{k'l'm'n'}\delta_{m'n'}\delta_{k'l'} + \ldots)\langle I|a_{i'}^\dagger a_{j'}^\dagger|J\rangle + \ldots \end{aligned}$$

Step 4c : Remove terms which known to be zero from index ranges.

Step 4d : Repeat process, but reorder to alternating order

$$\begin{aligned} \langle I|a_l a_i^\dagger a_m a_j^\dagger a_n a_k^\dagger|J\rangle \rightarrow \langle I|a_i^\dagger a_l a_j^\dagger a_m a_k^\dagger a_n|J\rangle \\ = \gamma^{IJ}_{ijklmn} \end{aligned}$$

Calculation of $\gamma^I_{ijk\ldots}$

Step 4e : Calculate $\gamma^I_{ijk\ldots}$ from a product of γ^I_{ij}, e.g.,

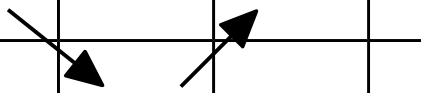
$$\gamma^I_{ijklmn} = \sum_{JKL} \langle I|a_i^\dagger a_j|K\rangle\langle K|a_k^\dagger a_l|L\rangle\langle L|a_m^\dagger a_n|J\rangle c_J$$

- Alternating order facilitates merging of terms and has some useful qualities.

Spin transition pathways

- Representation in terms of individual transitions can help with application of symmetry :

Spin sector	$\{\|J\rangle\}$	$\{\|K\rangle\}$	$\{\|L\rangle\}$	$\{\|I\rangle\}$
$[7\alpha 0\beta]$				
$[6\alpha 1\beta]$				
$[5\alpha 2\beta]$				
$[4\alpha 3\beta]$				
$[3\alpha 4\beta]$				
$[2\alpha 5\beta]$				
$[1\alpha 6\beta]$				
$[0\alpha 7\beta]$				



$$|J\rangle \in [4\alpha 3\beta]$$

$$\gamma_{\alpha\beta}^{KL} = \langle K|a_\alpha^\dagger a_\beta|L\rangle$$

$$\gamma_{\alpha\beta}^{IK}\gamma_{\alpha\beta}^{KL}\gamma_{\beta\alpha}^{LJ}$$

Spin transition pathways

- “Forwards” and “backwards” transitions are connected by time reversal

Spin sector	$\{\|J\rangle\}$	$\{\|K\rangle\}$	$\{\|L\rangle\}$	$\{\|I\rangle\}$
$[7\alpha 0\beta]$				
$[6\alpha 1\beta]$				
$[5\alpha 2\beta]$				
$[4\alpha 3\beta]$				
$[3\alpha 4\beta]$				
$[2\alpha 5\beta]$				
$[1\alpha 6\beta]$				
$[0\alpha 7\beta]$				

$$|J\rangle \in [4\alpha 3\beta]$$

$$\gamma^{KL}_{\beta\alpha} = \langle K|a^{\dagger}_{\beta}a_{\alpha}|L\rangle$$

$$\gamma^{JK}_{\beta\alpha}\gamma^{KL}_{\beta\alpha}\gamma^{LI}_{\alpha\beta}$$

$$\gamma^{KL}_{\beta\alpha} = (\gamma^{LK}_{\alpha\beta})^{*}$$

Spin transition pathways

- Application of spin constraints is more straightforward

Spin sector	$\{\|J\rangle\}$	$\{\|K\rangle\}$	$\{\|L\rangle\}$	$\{\|I\rangle\}$
$[7\alpha 0\beta]$				
$[6\alpha 1\beta]$				
$[5\alpha 2\beta]$				
$[4\alpha 3\beta]$				
$[3\alpha 4\beta]$				
$[2\alpha 5\beta]$				
$[1\alpha 6\beta]$				
$[0\alpha 7\beta]$				

$$|J\rangle \in [0\alpha 7\beta]$$

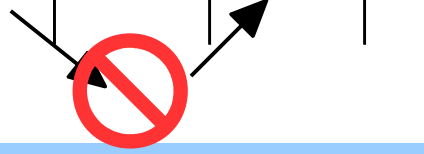
$$\rightarrow \gamma^{KJ}_{\beta\alpha} = 0$$

$$\rightarrow \gamma^{IK}_{\alpha\beta}\gamma^{KL}_{\alpha\beta}\gamma^{LJ}_{\beta\alpha} = 0$$

Spin transition pathways

- Application of spin constraints is more straightforward

Spin sector	$\{\|J\rangle\}$	$\{\|K\rangle\}$	$\{\|L\rangle\}$	$\{\|I\rangle\}$
$[7\alpha 0\beta]$				
$[6\alpha 1\beta]$				
$[5\alpha 2\beta]$				
$[4\alpha 3\beta]$				
$[3\alpha 4\beta]$				
$[2\alpha 5\beta]$				
$[1\alpha 6\beta]$				
$[0\alpha 7\beta]$				



$$|J\rangle \in [0\alpha 7\beta]$$

$$\rightarrow \gamma^{KJ}_{\beta\alpha} = 0$$

$$\rightarrow \gamma^{IK}_{\alpha\beta}\gamma^{KL}_{\alpha\beta}\gamma^{LJ}_{\beta\alpha} = 0$$

Calculation of $\gamma^I_{ijk\ldots}$

Step 4 : Calculate $\gamma^I_{ijk\ldots}$ from a product of γ^I_{ij}, e.g.,

$$\gamma^I_{ijklmn} = \sum_{JKL} \langle I|a_i^\dagger a_j|K\rangle\langle K|a_k^\dagger a_l|L\rangle\langle L|a_m^\dagger a_n|J\rangle c_J$$

- Alternating order enables merging of terms and some useful properties.
- All indexes must be active, but can be α or β.
- The program is designed so that arbitrary range constraints, not necessarily involving spin, can be placed upon the indexes.

Method outline

Step 1 : Read in tensors corresponding to operators represented in the molecular orbital basis.

Step 2 : Use constraints and symmetry conditions to determine unique blocks.

Step 3 : Determine all possible contractions, $A_{ijk\ldots}$, of these tensors.

Step 4 : Calculate γ^{IJ} matrices for the relevant spin sector.

Step 5 : Contract γ^{IJ} matrices with the relevant $A_{ijk\ldots}$.

Guiding principal : Relativistic and non-relativistic cases should be treated using the same basic algorithm.

→ *Relativistic wavefunctions just have twice as many blocks (from spin), and extra symmetry relations.*

Contraction of γ^{I}_{ijklmn} and A_{ijklmn}

Step 5 : Perform contraction in two steps:

$$\sum_{\sigma_3}\sum_{ijklmn}\gamma^{IJ\sigma_3}_{ij,kl,mn}A^{\sigma_3}_{ij,kl,mn}c_J=\sum_{\sigma_1}\sum_{ij}\sum_{K}\gamma^{IK\sigma_1}_{ij}\tilde{\gamma}^{K\sigma_1}_{ij}$$

- Where

$$\tilde{\gamma}^{K\sigma_1}_{ij}=\sum_{\sigma_2}\sum_{LJ}\sum_{klmn}\gamma^{KJ\sigma_2}_{kl,mn}A'^{\sigma_1\otimes\sigma_2}_{ij,kl,mn}$$

and

$$\sigma_3=\sigma_1\cup\sigma_1\cup\sigma_1=\{s_1s_2s_3s_4s_5s_6\}$$

- This step should be comparable in speed to the most expensive step of the current algorithm.

Calculation of $\gamma^I_{ijk\ldots}$

- Use CAR relations to minimize the number of spin flipping terms, and ensure they are always grouped on the right, i.e.,

$$
\begin{array}{llll}
(\alpha\alpha)(\alpha\alpha)(\alpha\alpha) & (\beta\beta)(\beta\beta)(\beta\beta) & (\beta\alpha)(\beta\alpha)(\beta\alpha) & (\alpha\beta)(\alpha\beta)(\alpha\beta) \\
(\alpha\alpha)(\alpha\alpha)(\beta\alpha) & (\beta\beta)(\beta\beta)(\beta\alpha) & & \\
(\alpha\alpha)(\beta\alpha)(\beta\alpha) & (\beta\beta)(\beta\beta)(\alpha\beta) & & \\
(\alpha\alpha)(\alpha\alpha)(\alpha\beta) & (\beta\beta)(\alpha\beta)(\alpha\beta) & & \\
(\alpha\alpha)(\alpha\beta)(\alpha\beta) & (\beta\beta)(\beta\alpha)(\beta\alpha) & & \\
(\alpha\alpha)(\beta\beta)(\beta\beta) & & & \\
(\alpha\alpha)(\beta\beta)(\beta\alpha) & & & \\
(\alpha\alpha)(\beta\beta)(\alpha\beta) & & &
\end{array}
$$

$$(\alpha\beta) = \langle I | a^\dagger_{i\in\{\alpha\}} a^\dagger_{i\in\{\beta\}} | J \rangle$$

- Guarantees that the number of contractions to be performed in step 5 is at worst 16 (typically far fewer due to symmetry).

Method outline

Step 1 : Read in tensors corresponding to operators represented in the molecular orbital basis.

Step 2 : Use constraints and symmetry conditions to determine unique blocks.

Step 3 : Determine all possible contractions, $A_{ijk\ldots}$, of these tensors.

Step 4 : Calculate γ^{IJ} matrices for the relevant spin sector.

Step 5 : Contract γ^{IJ} matrices with the relevant $A_{ijk\ldots}$.

Guiding principal : Relativistic and non-relativistic cases should be treated using the same basic algorithm.

→ *Relativistic wavefunctions just have twice as many blocks (from spin), and extra symmetry relations.*