# 1 Gamma generator

Consider a term

$$\sum_{ijkl}\sum_{mnop}\sum_{IJ}\langle I|i^\dagger j^\dagger klm^\dagger n^\dagger op|J\rangle c_I^M c_J^N Y_{ijkl}^{ML} Z_{mnop}^{NP}, \tag{1}$$

where $I$ and $J$ each denote a Slater determinant, and $L, M, N, P$ are state indexes. $\mathbf{Y}^{ML}$ and $\mathbf{Z}^{NP}$ are representations of operators and $\hat{Y}$ and $\hat{Z}$ in the molecular orbital basis. For the time being the state dependence of the operator representations will be ignored, but has important consequences for exploitation of symmetry, and will be discussed at length later.

The procedure most commonly employed by code generators is to use Wick's theorem to rearrange expressions, such as (1), into a series of normal ordered terms, e.g.,

$$\begin{aligned}
&= \sum_{stuvwxyz}\sum_{IJ}\langle I|s^\dagger t^\dagger u^\dagger v^\dagger wzyx|J\rangle c_I c_J^* A_{stuvwxyz},\\
&+ \sum_{stuwxy}\sum_{IJ}\langle I|s^\dagger t^\dagger u^\dagger wzy|J\rangle c_I c_J^* A_{stuwxy},\\
&+ \sum_{stwx}\sum_{mnop}\sum_{IJ}\langle I|s^\dagger t^\dagger wz|J\rangle c_I c_J^* A_{stwx},\\
&+ \sum_{sw}\sum_{mnop}\sum_{IJ}\langle I|s^\dagger w|J\rangle c_I c_J^* A_{sw}\\
&+ \sum_{sw}\sum_{mnop} A.
\end{aligned}$$

Where the tensors $A_{ijkl\ldots}$ are formed by performing the contractions between and re-orderings of the indexes of tensors on $Y$ and $Z$ as determined from the commutation relations of the creation and annihilation operations, e.g.,

$$A_{stwx} = \sum_{r}^{R} \hat{\wp}_r \sum_{\{u,y\}}^{c1} \delta_{uy} \sum_{\{v,z\}}^{c2} \delta_{vz} Y_{ijkl} Z_{mnop},$$

where the $c1$ and $c2$ are sets of pairs of indexes to be contracted, e.g.,

$$c2 = \{\{i,k\}\},\{j,k\}\},\{l,m\}\}.....\} \tag{2}$$

and $\hat{\wp}_r$ which transforms the ordered set of indexes $\{u,v,s,t,w,x,y,z\}$ into some permutation, $r \in R$, of the ordered set of indexes $\{i,j,k,l,m,n,o,p\}$, whilst also acting to multiply the result of the summation by an appropriate factor.

A major advantage of this is that it enables constraints to be on the values over which the

molecular orbital indexes range; all indexes on annihilation operators must correspond to orbitals occupied in $|J\rangle$, and creation operators must correspond to orbitals occupied in $|I\rangle$. This is a consequence of how determinants are constructed in active space based methods. A generic determinant can be written corresponding to the case where the, molecular orbitals, $\{\psi_{x_i}\}$, with indexes $X = \{x_0, ..., x_n\}$ are occupied can be written

$$\Psi_K = \sum_{\zeta}^{\zeta \in S_X} \epsilon_\zeta \bigotimes_{i=0}^{n} \psi_{x_i}, \tag{3}$$

where $X$ is a set of molecular orbital indexes $\{x_0, ..., x_n\}$, $\epsilon$ is the nth rank Levi-Cevita tensor, $\zeta$ is a member of the symmetric group $S_X$, consisting of all possible permutations on the set of indexes $X$.

In CI based methods it is common to use the fact that each distinct ordered set, $X$, of indexes corresponds to a different determinant (given all sets being compared have their indexes ordered according to some standard set of rules). In active space based methods the set of determinants used to define the wavefunction is constrained to only those determinants which corresponding $X \in F^{CAS}$, where

$$F^{CAS} = \{X | (X = X^c \cup X^a) \wedge (X^a \subset P) \wedge (\text{card}(X) = n_{act})\}$$

where $P$ is the list of all "active" molecular orbital indexes, $X^a$ is a subset of active orbitals occupied in the determinant defined by $X$, and where $X^c$ is the list of closed orbital indexes, all of which are occupied in all determinants defined from members of $F^{CAS}$.

Unfortunately, $F^{CAS}$ can get very large, leading to computational difficulties. One way of trying to manage this is to decompose the set of active orbital indexes, $P$, up into $n_P$ subsets;

$$P = \bigcup_{j=0}^{n_P} P_j, \tag{4}$$

and then splitting $F^{CAS}$ up into subsets, $F_{o_1,...,o_{n_p}}$, based on the number, $o_j$, of orbitals in each subspace, $P_j$, which are occupied :

$$F_{o_1,...,o_{n_p}} = \left\{ X = X_c \cup \left[ \bigcup_{j=0}^{n_p} X_j \right] \middle| (X_j \subset P_j) \wedge \left( \sum_{j}^{n_p} o_j = n_{act} \right) \right\}. \tag{5}$$

Here $X_j$ is the subset of $P_j$ with cardinality $o_j$. The second constraint is just the requirement that the sum of all the occupancy numbers, $o_j$, associated with each of the different subspaces, $P_j$, adds up to the total number of occupied electrons.

To illustrate this consider the case where $P$ is split into two subspaces;

$$P = P_\mu \cup P_\nu. \tag{6}$$

For the case where

$$\text{card}(P_\mu) = 2, \quad \text{card}(P_\nu) = 4, \quad \text{and} \quad n_{act} = 3$$

The space, $F^{CAS}$, can then be split up according to how many $\gamma$ and $\nu$ electrons are occupied in a given element;

$$F^{CAS} = F^{CAS}_{3,0} \cup F^{CAS}_{2,1} \cup F^{CAS}_{1,2}$$

$$F^{CAS}_{3,0} = \{X| \wedge (X^a_\mu \subset P_{mu}) \wedge (\text{card}(X)_\mu = n_{act}\}$$

$$F^{CAS}_{2,1} = \{X| \wedge (X^a_\mu \subset P_{mu}) \wedge (\text{card}(X)_\mu = n_{act} - 1\} \wedge (X^a_\nu \subset P_{nu}) \wedge (\text{card}(X)_\nu = 1\}$$

$$F^{CAS}_{1,2} = \{X| \wedge (X^a_\mu \subset P_{\mu}) \wedge (\text{card}(X)_\mu = n_{act} - 2\} \wedge (X^a_\nu \subset P_{nu}) \wedge (\text{card}(X)_\nu = 2\}$$

$$X = X^c \cup X^a_\mu \cup X^a_\nu.$$

I shall hereafter refer to these subspaces of $F^{CAS}$ as ci-sectors (and in specific cases, spin-sectors). Decomposing the active space in this manner has a number of benefits, as the different ci-sectors may not interact, or have their interaction limited in some way. The most prominent example of this is the non-interaction of ci-sectors with different numbers of occupied $\alpha$ and $\beta$ orbitals in the non-relativistic framework. However, even if they do, the decomposition of the active space makes it is possible to handle this interaction a piece wise manner, i.e., deal with interactions between two ci-sectors at a time. *Note : Please correct me if I am wrong, but I interpret relativistic CI as a kind of A.S.D., albeit with very different rules governing the interactions between the different sectors. I realize the implementations and approaches to optimization are very different.*

One major advantage is that decomposition of the active-space into different components facilitates the block wise decomposition of the reduced density matrices (RDMs), as a block of an RDM can being defined by the CI-sectors to which the Bra and Ket belong. Using the above decomposition of the active space as an example;

$$\sum_{J}^{\in\mathcal{F}_{all}} \sum_{I}^{\in\mathcal{F}_{all}} \langle I|i^\dagger j^\dagger m^\dagger n^\dagger klop|J\rangle c^M_I c^N_J$$

$$\sum_{\mathcal{F}_{sub}}^{\in\mathcal{F}_{all}} \sum_{\mathcal{F}_{sub}}^{\in\mathcal{F}_{all}} \sum_{I}^{\in\mathcal{F}_{sub}} \sum_{J}^{\in\mathcal{F}_{sub}} \langle I|i^\dagger j^\dagger m^\dagger n^\dagger klop|J\rangle c^M_I c^N_J,$$

where $\mathcal{F}_{sub}$ is one of the sub-spaces into which the total Fock space $\mathcal{F}_{all}$ has been decomposed.

Unfortunately, this decomposition brings with it a number of complexities which are not present for undecomposed active spaces. However, use of such decomposition techniques can lead to significant gains in computational efficiency, particularly for methods which much handle multi-reference, relativistic wavefunctions.

Allowance for the fact that the Bra and Ket may belong to different CI-sectors, the blockwise decomposition of the molecular orbital tensors, and the exploitation of symmetry which exists within and between different blocks *and* different CI-sectors, are the distinguishing features of the program.

## 1.1 Blockwise handling of terms

That different the Bra and Ket may have different CI-sectors has important ramifications for the program structure, however, before discussing these in depth it is necessary to discuss the basics of the task list construction.

In the previous section we discussed the construction of the task list used for calculating contractions between multiple different tensors, e.g., the task list for obtaining

$$A_{kwzm} = [\tilde{B}\tilde{C}\tilde{D}]^{(jy,lx,no,pi)}_{kwzm} = \sum_{jy}\sum_{no}\sum_{lx}\sum_{ip} B_{ijkl}C_{mnop}D_{wxyz}\delta_{jy}\delta_{no}\delta_{lx}\delta_{ip}. \tag{7}$$

The construction of this "contraction task list" (referred to as *A_compute_list* in the code, is effectively independent from the task list to be discussed now, which concerns determination of which contracted tensors, e.g., which $A_{kwzm}$, that need to be calculated in order to obtain the expectation value or derivative.

As stated in the program overview, the total equation (or expression) to be solved (or evaluated), is broken down into a number of *Terms*, e.g.,

$$\langle M|\hat{Y}\hat{Z}|N\rangle,$$

which using second quantization as

$$\sum_{ijkl}\sum_{mnop}\sum_{IJ}\langle I|i^\dagger j^\dagger klm^\dagger n^\dagger op|J\rangle c_I^M c_J^N Y_{ijkl}^{ML} Z_{mnop}^{NP}. \tag{8}$$

In (8) the sum over the orbital indexes $\{i,j,k,l\}$ and $\{m,n,o,p\}$ runs over all indexes specified by tensors $\mathbf{Y}$ and $\mathbf{Z}$ respectively. However, we can rewrite the above in terms of summations

over the CI-sectors and blocks into which these tensors may be decomposed;

$$\sum_{B^Y}\sum_{B^Z}\sum_{ijkl}^{B^Y}\sum_{mnop}^{B^Z}\sum_{IJ}\langle I|i^\dagger j^\dagger klm^\dagger n^\dagger op|J\rangle c_I^M c_J^N Y_{ijkl}^{ML} Z_{mnop}^{NP}. \tag{9}$$

Here $B^Y$ and $B^Z$ are blocks of $\mathbf{Y}$ and $\mathbf{Z}$ which define the ranges over which the summations of the molecular orbital indexes are to occur.

Many of the contributions from these various tensor blocks will vanish, but exactly which will vanish is dependent on the CI-sector to which $|I\rangle$ and $|J\rangle$ belong. For example, if

$$|I\rangle \in \mathcal{F}_{\mu=2,\nu=1} \qquad \text{and} \qquad |J\rangle \in \mathcal{F}_{\mu=3,\nu=0},$$

then the cumulative action of the creation and annihilation operators within the BraKet must be to destroy one electron in range $r^\mu$, and create one in $r^\nu$. All terms corresponding to blocks which do not meet this criterion can immediately be discarded. This idea can be taken further: If the expression is rearranged into normal order,

$$\sum_{B^Y}\sum_{B^Z}\sum_{ijkl}^{B^Y}\sum_{mnop}^{B^Z}\sum_{IJ}\langle I|i^\dagger j^\dagger m^\dagger n^\dagger klop|J\rangle c_I^M c_J^N Y_{ijkl}^{ML} Z_{mnop}^{NP}. \tag{10}$$

then not only do we retain the above constraint, but it is also known that terms corresponding to a combination of blocks, $B^Y$ and $B^Z$, where *any* of the indexes $k$, $l$, $o$, $p$ are in range $\nu$, will vanish. An analogous constraint involving constraints on the ranges of the creation operators is obtained by putting the expression into anti-normal ordering;

$$\sum_{B^Y}\sum_{B^Z}\sum_{ijkl}^{B^Y}\sum_{mnop}^{B^Z}\sum_{IJ}\langle I|klopi^\dagger j^\dagger m^\dagger n^\dagger|J\rangle c_I^M c_J^N Y_{ijkl}^{ML} Z_{mnop}^{NP}. \tag{11}$$

By applying this procedure we ensure that only indexes within the active space are left within the BraKet, which is computationally significant as the dimension of the active orbital subrange, $r^a$, is typically much smaller than the other subranges, such as those corresponding to core, $r^c$, or virtual, $r_v$, orbitals.

Whilst this rearrangement will generate new terms in accordance with the commutation relations of the creation and annihilation operators, all of these new terms will be of lower rank than the original. This fact, combined with the range constraints and elimination of terms means that the increase number of computational tasks to be performed is generally well worth it.

Perhaps most importantly of all this means calculations can be done in a piece-wise fashion,

treating only one combination of CI-sectors at a time. This is particularly useful when calculating expressions requiring reduced density matrix derivatives,

$$\Gamma^{I}_{ijklmn} \sum_{J} \langle I|i^{\dagger}j^{\dagger}m^{\dagger}n^{\dagger}opkl|J\rangle c_{J}, \tag{12}$$

which due to the lack of a summation over one of the CI-indexes, $I$, can prove extremely large. could prove highly problematic for implementations of energy derivative expressions in the relativistic framework; the treatment of $\alpha$ and $\beta$ electrons effectively doubles the size of the active space. However, the block decomposition of molecular orbital tensors, the decomposition of the active space, and the seperation of blocks into a real and imaginary components, ensures that the peak memory requirements of the calculations need not exceed those for similar expressions in a non-relativstic framework.

A disadvantage is that the algebraic operations of rearranging the tensors must be performed for every possible block, of which there can in principal be millions[1]. Fortunately, there are a number of ways to mitigate this, but it is not so trivial an issue that it can be completely ignored.

The reordering sequence specified above; normal ordering followed by anti-normal ordering, encapsulates the most important principals used by the algebraic manipulator. However, more efficient task lists can be generated by following these two initial reorderings with others, the structure of which are determined by the properties of the block, the term being calculated[2]. In order to understand why this approach is taken, and how these reorderings are chosen and generated, it is necessary to discuss the implementation of symmetry.

# 2 Handling of block and spin-sector symmetry

One of the defining features of the program is that symmetry is incorporated at the stage of algebraic manipulation, and precedes the task list construction, i.e., it is only after all expressions have been rewritten to involve some minimal set of distinct terms that the task list is generated. Thus if two terms are equivalent, one will of these terms will never appear at any stage of the final task list. This is primarily to facilitate the identification of terms which can be merged or reused (to save memory), and terms which can be completely seperated (to facilitate parallization).

Whilst the previous section discussed how symmetry was applied in execution and representation of the contraction task list, it did not discuss how symmetry was used to decide what

[1] If each range is divided in to six, and we have ten indexes, $6^{10} = 60466176$.

[2] For example, terms involving CI-derivatives use different reorderings to terms involving molecular orbital tensor derivatives.

quantities that task list was attempting to calculated. The focus of this section is the latter of these two points.

Broadly speaking, the algorithm should function such that given the following equivalences:

$$\hat{T}^{KM} = \hat{T}^{KP} \ , \quad \hat{T}^{-1} = T^{\dagger} \ , \text{ and } \quad \mathbf{T}^{KM}_{b1} = \mathbf{T}^{KP}_{b2}$$

$$\mathbf{H}_{b3} = \mathbf{H}_{b4}$$

$$\mathbf{c}^{P}_{K} = -\boldsymbol{\sigma_{y,2\times 2}}\mathbf{c}^{M*}_{J}, |K\rangle = |\overline{J}\rangle, \quad \text{where} \quad \boldsymbol{\sigma_{y,2\times 2}} = \begin{bmatrix} \boldsymbol{\sigma_y} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma_y}, \end{bmatrix}$$

any contribution written using the terms on the left hand side (LHS) of the above can be rewritten as one involving the terms on the right hand side (RHS) of the above. For example,

$$\sum_{ijkl}^{b1}\sum_{mnop}^{b2}\langle K|\hat{i}^{\dagger}\hat{j}^{\dagger}\hat{k}\hat{l}\hat{m}^{\dagger}\hat{n}^{\dagger}\hat{o}\hat{p}|K\rangle \mathbf{c}^{P}_{K}\mathbf{c}^{M}_{I}T^{KM,-1,b1}_{ijkl}H^{b3}_{mnop}$$

is re-expressed as this term

$$\sum_{ijkl}^{b1}\sum_{mnop}^{b3}\langle I|\hat{m}^{\dagger}\hat{n}^{\dagger}\hat{o}\hat{p}\hat{\overline{i}}^{\dagger}\hat{\overline{j}}^{\dagger}\hat{\overline{k}}\hat{\overline{l}}|\overline{J}\rangle \mathbf{c}^{M*}_{J}\mathbf{c}^{M}_{I}H^{b4}_{mnop}T^{KM,b2}_{lkji}$$

plus some sequence of transformations, and potentially a number of lower order terms which will arise due to the switching of the order of $\hat{H}$ and $\hat{T}$. Here the notations $\overline{i}$ is the index corresponding to the Kramer's pair of orbital $i$.

The key thing to note in the above is that the transformations of the blocks range limits on the summations are not transformed, whilst those on the molecular orbital tensors are. Furthermore, the indexes inside the BraKet are not transformed with the representations of the molecular orbital tensors, but are transformed due to the time reversal symmetry operation resulting from the substitution $|K\rangle = |\overline{I}\rangle$, and the hermitian conjugation operation applied to $\hat{T}$.

This is because transformations of the operator, such as inversion and time-reversal, are fundamentally different to transformations of to the representation of an operator in the molecular orbital basis, e.g., mapping $T^{b1}_{ijkl} \rightarrow T^{b2}_{klij}$. Transformations to the operator itself will impact the ranges on which the creation and annihilation operators used to define that operator act, e.g., in the above case time-reversal alters the ranges on which the creation and annihilation operators act, and hence alters the ranges within the BraKet. Conversely, symmetries arising from properties of the molecular orbital indexes on the molecular orbital tensors do not influence the ranges on which the creation and annihilation operators act, as these are symmetries

arising from the molecular orbital basis chosen to represent the operator.

Clearly, there is a very close link between the two, but they are distinct. By way of example consider an operator, $\hat{W}$, which is split into two blocks

$$b1 = \{\alpha, \alpha, \alpha, \alpha\} \qquad \text{and} \qquad b2 = \{\beta, \beta, \beta, \beta\}$$

ranges on indexes. Suppose that ranges $r_\alpha$ and $r_\beta$ have equal extent, and that the representation of $\hat{W}$ within each of these blocks is equivalent, i.e.,

$$W^{b1}_{ijkl} = W^{b2}_{ijkl}. \tag{13}$$

Now consider a 2-electron wavefunction, formed from a linear combination of determinants, in each of which two alpha orbitals are occupied

$$|M\rangle = \sum_I c_I^M |I(n_\alpha = 2, n_\beta = 0\rangle. \tag{14}$$

We can now write

$$\langle M|\hat{W}|M\rangle = \langle M|\hat{W}^{b1}|M\rangle + \langle M|\hat{W}^{b2}|M\rangle$$

$$= \sum_{ijkl}^{b1}\sum_{IJ}\langle I|\hat{i}^\dagger\hat{j}^\dagger\hat{k}\hat{l}|J\rangle c_I c_J W^{b1}_{ijkl} + \sum_{ijkl}^{b2}\sum_{IJ}\langle I|\hat{i}^\dagger\hat{j}^\dagger\hat{k}\hat{l}|J\rangle c_I c_J W^{b2}_{ijkl},$$

which then using $W^{b1}_{ijkl} = W^{b2}_{ijkl}$ leads to

$$= \sum_{ijkl}^{b1}\sum_{IJ}\langle I|\hat{i}^\dagger\hat{j}^\dagger\hat{k}\hat{l}|J\rangle c_I c_J W^{b1}_{ijkl} + \sum_{ijkl}^{b2}\sum_{IJ}\langle I|\hat{i}^\dagger\hat{j}^\dagger\hat{k}\hat{l}|J\rangle c_I c_J W^{b1}_{ijkl}.$$

However, it is clear from the definition of $|M\rangle$ that the second term must always be zero, as there are no $\beta$ electrons to destroy, hence,

$$\neq 2\sum_{ijkl}^{b1}\sum_{IJ}\langle I|\hat{i}^\dagger\hat{j}^\dagger\hat{k}\hat{l}|J\rangle c_I c_J W^{b1}_{ijkl}.$$

This means that whilst we can transform all occurrences of $\mathbf{W}^{b2}$ into occurrences of $\mathbf{W}^{b1}$, we are typically unable to apply the same transformation to the block ranges and the molecular orbital indexes which appear withing the braket. Transformations of these indexes only arise when performing transformations on the physical action of the operator, such as time-reversal, or inversion.

It is for this reason that symmetries are not dealt with by using a single $i < j < k < l$ type rule, but instead are subdivided into three distinct types, which can be combined to guide construc-

tion of the task list. These three basic symmetry types are : operator symmetry, representation (or block) symmetry, and CI-sector symmetry.

Operator symmetry refers to transformations on the action of the operator, and equivalences between operators which act on different combinations of states (dictated by time reversal, or potentially spatial symmetry). For example,

$$\hat{T}^{LM} = \hat{T}^{\overline{LM}}_{ijkl} \tag{15}$$

where $\overline{L}$ is the index referring to the Kramers pair of the state with index $L$. Another kind of operator symmetry is

$$T^{LM,b}_{ijkl} = T^{\overline{LM},\bar{b}}_{ijkl} \tag{16}$$

Representation or block symmetries are relations such as

$$\mathbf{T}^{LM,b1}_{ijkl} = \rho \mathbf{T}^{LM,b2}_{klij} \tag{17}$$

where $b2$ is some index block which may be obtained by performing a permutation operation of the indexes of $b1$, and $\rho$ is some multiplicative factor. Another kind of common block symmetry is

$$\mathbf{T}^{LM,b}_{ijkl} = \rho \mathbf{T}^{\overline{LM},\bar{b}}_{ijkl} * \tag{18}$$

$\bar{b}$ is the range block defined from the index ranges of the which are Kramers pairs of the molecular orbital index ranges which define block $b$. Similarly, $\overline{L}$ and $\overline{M}$ are the indexes of the states which are Kramers pairs of $L$ and $M$ respectively.

Finally, and example of CI-sector symmetry would be

$$\mathbf{c}^{P}_{K} = -\boldsymbol{\sigma}_{\boldsymbol{y},\mathbf{2}\times\mathbf{2}} \mathbf{c}^{M*}_{J}, |K\rangle = |\overline{J}\rangle, \quad \text{where} \quad \boldsymbol{\sigma}_{\boldsymbol{y},\mathbf{2}\times\mathbf{2}} = \begin{bmatrix} \boldsymbol{\sigma}_{\boldsymbol{y}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_{\boldsymbol{y}} \end{bmatrix}$$

This is just time reversal symmetry.

Initially, the state dependence of the operators seems perculiar; the molecular orbitals themselves do not differ between states, hence the state dependence of the representations suggests that it is not possible to interpret these operators as describing interactions between 1, 2, or n-electrons. In fact, such state dependent operators do not correspond to physical interactions per se, but are instead a tool used to aid in the description of representation of a perturbed state in terms via interaction of a number of unperturbed states. An important consequence of this is that the symmetry of the operator representations is determined by the form of the equations from which they are obtained, rather than from consideration of the form of operators

themselves.