Grassmann

Yuri Alexandre Aoto

Contents

Preface

Part I

Using the program

Part II

Theory and implementation

Chapter 1

Optimisation of distance in the Grassmannian to an external wave function

YURI ALEXANDRE AOTO

History:

24 mar 2019 - Start

31 oct 2019 - Adding Absil method

30 mar 2020 - Adding procedure for CISD wave function

1.1   Notation

TODO: (??) Change position of spin and irrep in the notation? $(n_\sigma^\Gamma \to n_\Gamma^\sigma)$

$\mathcal{H}$	The electronic Hamiltonian
g	Number of irreducible representations (short: irrep) in the point group
Γ	An irreducible representation, or an index that runs over all irreps
σ	spin (α or β)
n_σ^Γ	The number of electrons with spin σ and irrep Γ
	(This might refer to a particular configuration, or to the reference
	determinant. The context should make it clear.)
n_σ	The number of electrons with spin σ
	$n_\sigma = \sum_{\Gamma=1}^g n_\sigma^\Gamma$
n	The total number of electrons
	$n = n_\alpha + n_\beta$
$\mathcal{V}$	The space of spatial orbitals (one-electron wave functions)
	(The full space of one-electron wave functions is assumed to be a direct
	sum of two such spaces, the first associated to α-spin and the
	second to β-spin electrons. A subscript might be used, but the
	space is the same: $\mathcal{V}_\alpha = \mathcal{V}_\beta = \mathcal{V}$)
$\mathcal{V}^\Gamma$	The space of orbitals of irreducible representation (short: irrep) Γ
	($\mathcal{V}$ is the direct sum of all such $\mathcal{V}^\Gamma$: $\mathcal{V} = \bigoplus_{\Gamma=1}^g \mathcal{V}^\Gamma$)
I, J	Multi-indices, usually of size n, n_σ, or n_σ^Γ, depending on the context.
K	The dimension of $\mathcal{V}$: $K = \dim \mathcal{V}$
K_Γ	The dimension of $\mathcal{V}^\Gamma$: $K_\Gamma = \dim \mathcal{V}^\Gamma$
	$\sum_\Gamma K_\Gamma = K$
i,j,k,l	indices for occupied and orbitals, that are correlated in the wave function
	(the context should make clear the reference determinant and which wave
	function they refer to)
a,b,c,d	indices for virtual orbitals
	(the context should make clear the reference determinant)
p,r	indices running over all K orbitals
q,s	indices running over all n occupied orbitals (correlated and not correlated)
$\sigma(p)$	spin of orbital p
$\mathcal{H}$	The complete Hilbert space of quantum states
	(within finite basis set approximation)
	$\mathcal{H} = \bigwedge^n (\mathcal{V}_\alpha \oplus \mathcal{V}_\beta)$
Ψ	General elements of $\mathcal{H}$
Φ	Elements of $Gr(n, 2K) \in \mathcal{H}$, that is, they can be expressed
	as a Slater determinant.
ϕ, ψ	Elements of $\mathcal{V}$ (orbitals)

1.2 Introduction

Let

$$|\Psi_{\text{ext}}\rangle = \sum_I C_I |\Phi_I\rangle \tag{1.1}$$

be a n-electron normalised CI like wave function represented in the orbital basis

$$|\phi_I\rangle = \phi_{I_1} \wedge \phi_{I_2} \wedge \ldots \wedge \phi_{I_n} , \tag{1.2}$$

where

$$\{\phi_p\}_{p=1}^K \tag{1.3}$$

is an orthonormal basis for the space of one-electron wave functions and I is an ordered multi-index. We want to find $|\Phi\rangle \in Gr$ such that $|\langle\Psi_{\text{FCI}}|\Phi\rangle|$ is maximum, where Gr is the image of the Grassmannian in the space of the n-electron wave functions.

Recall that

$$D(\Psi_1, \Psi_2) = \sqrt{2}\sqrt{1 - |\langle\Psi_1|\Psi_2\rangle|} \tag{1.4}$$

is a metric in $\mathbb{P}\mathscr{H}$.[] Also, Gr is the set of all elements of $\mathscr{H}$ that can be written as a single Slater determinant (a decomposable element) for some orthonormal basis of $\mathcal{V}$. For this reason, such $|\Phi\rangle$ minimises the distance D to the wave function $|\Psi_{\text{ext}}\rangle$ and it will be denoted $|\Phi_{\text{minD}}\rangle$. The subscript "ext" refers to *external*, because we assume that $|\Psi_{\text{ext}}\rangle$ is external to the Grassmannian.

We have developed and implemented three algorithms to calculate $|\Phi_{\text{minD}}\rangle$. The first uses the whole structure of a Full-CI wave function (that is, the full Hilbert space), because the external wave function has to be transformed to the MO basis of $|\Phi_{\text{minD}}\rangle$. Hence, even if we start with a CISD wave function for example, when we update $|\Phi_{\text{minD}}\rangle$ the orbitals change, and the external wave function is no longer just a CISD over that reference. The second and third use the Newton method adapted to the Grassmannian, as proposed by Absil *et al.* []. The second is a general implementation that works for, in principle, any wave function. The third is specific for closed shell restricted CISD wave functions.

For these last two cases spin adaption is implemented. However, for these last two cases we have algorithms that find critical points, but do not test for tru maximum

In any case, this document is intended to give the full formulas that are implemented in the present code. The detailed derivation of these expressions are not given here.

1.3   **Algorithm 1**

Using orbital rotations and the structure of a FCI wave function

To find $|\Phi_{\mathrm{minD}}\rangle$, it is equivalent to find the orbitals (namely, a basis of $\mathcal{V}$) such that

$$|\Psi_{\mathrm{FCI}}\rangle = C_0 |\Phi_{\mathrm{minD}}\rangle + \sum_{\substack{I \\ I \neq \{1,2,\dots,n\}}} C_I |\Phi_I\rangle \,, \tag{1.5}$$

and $|C_0|$ is maximum (over all possible coefficients in all possible basis), since $|C_0| = |\langle\Psi_{\mathrm{FCI}}|\Phi_{\mathrm{minD}}\rangle|$.

1.3.1   **Parametrisation by orbital rotations**

We parametrise Gr by the orbital rotations as[]

$$|\Phi\rangle = e^{-\hat{K}} |\Phi_0\rangle \,, \tag{1.6}$$

where

$$\hat{K} = \sum_{i,a} K_i^a (a_a^\dagger a_i - a_i^\dagger a_a) \tag{1.7}$$

$$= \sum_{i,a} K_i^a (a_i^a - a_a^i) \tag{1.8}$$

This parametrisation comes from the most general

$$\hat{K} = \sum_{p,r} K_r^p a_p^\dagger a_r \,, \tag{1.9}$$

but using that K_r^p is anti-symmetric (so that $e^{-\hat{K}}$ is orthogonal), and excluding rotations within the occupied or virtual spaces of $|\Phi_0\rangle$, that are redundant (do not alter the Slater determinant with $|\Phi_0\rangle$). For $\hat{K} = 0$ it is clear that $|\Phi\rangle = |\Phi_0\rangle$.

Let

$$f(K_i^a) = \left| \langle\Psi_{\mathrm{FCI}}|e^{-\hat{K}}|\Phi_0\rangle \right| \,, \tag{1.10}$$

where the argument K_i^a represent all the $n_\alpha(N - n_\alpha) + n_\beta(N - n_\beta)$ elements. Note that rotations that mix α and β orbitals are not considered. We will also assume that $\langle\Psi_{\mathrm{FCI}}|\Phi_0\rangle > 0$ and this remains true for all steps of our optimisation. If $\langle\Psi_{\mathrm{FCI}}|\Phi_0\rangle < 0$ we of course can change the phase of the wave function. If $\langle\Psi_{\mathrm{FCI}}|\Phi_0\rangle = 0$ for the first or any step of the optimisation, we likely started with a very poor initial guess.

spin and symmetry restricted

If the orbital rotations will preserve spin and symmetry, the operator $\hat{K}$ can be written as:

$$\hat{K} = \sum_{\Gamma} \sum_{(i,a)\in\Gamma} K_i^{a,\Gamma} (E_{ai}^\Gamma - E_{ia}^\Gamma), \tag{1.11}$$

where the singlet excitation operator is defined as:

$$E_{pr} = a_{p\alpha}^\dagger a_{r\alpha} + a_{p\beta}^\dagger a_{r\beta} \,. \tag{1.12}$$

See [**?**] for the details.

1.3.2 Jacobian and Hessian

We want to maximise f and we need its Jacobian and Hessian. The expressions at $\hat{K} = 0$ are given below. For the derivation of the expressions, see that hand notes.

$$\frac{\partial f(\hat{K} = 0)}{\partial K_i^a} = (-1)^{n_{\sigma(i)} - i + 1} C_i^a \tag{1.13}$$

$$\frac{\partial^2 f(\hat{K} = 0)}{\partial K_i^a \partial K_j^b} = \begin{cases} -C_0 & (i = j, a = b) \\ 0 & (i \neq j, a = b) \\ 0 & (i = j, a \neq b) \\ (-1)^{n_{\sigma(i)} + n_{\sigma(j)} - i - j} C_{ij}^{ab} & (\sigma(i) \neq \sigma(j)) \\ (-1)^{i+j+1} C_{ij}^{ab} & (\sigma(i) = \sigma(j), i < j, a < b) \\ (-1)^{i+j} C_{ij}^{ab} & (\sigma(i) = \sigma(j), i < j, a > b) \end{cases} \tag{1.14}$$

In these equations, C_i^a and C_{ij}^{ab} are the CI coefficients of the single and double excited determinants in the (normalised) wave function $|\Psi_{\text{FCI}}\rangle$. The canonical order of the orbitals is assumed to be "first all α, then all β".

Spin restricted, closed-shell reference and symmetry adapted

The above equations for the Hessian and the Jacobian are for a general wave function. Here we will show the corresponding equations for the case of a reference wave function, $|\Phi_0\rangle$ that is closed shell, restricted, and symmetry adapted. The external wave function $|\Psi_{\text{FCI}}\rangle$ will also be assumed to be restricted and symmetry adapted. In this conditions, the Jacobin and the Hessian are:

$$\frac{\partial f(\hat{K} = 0)}{\partial K_i^{a,\Gamma}} = -2(-1)^{i+n_\Gamma} C_i^{a,\Gamma} \tag{1.15}$$

$$\frac{\partial^2 f(\hat{K} = 0)}{\partial K_i^{a,\Gamma} \partial K_{i'}^{a'\Gamma'}} = \begin{cases} -2C_0 + 2\mathcal{C}_{ii}^{aa,\Gamma} & \Gamma = \Gamma', a = a', i = i' \\ 2\mathcal{C}_{i'i}^{a'a,\Gamma}(-1)^{i+i'} & \begin{cases} \Gamma = \Gamma'; \\ a = a' \text{ and } i \neq i' \text{ or} \\ a \neq a' \text{ and } i = i' \end{cases} \\ 2(-1)^{i+i'} \left(\mathcal{C}_{i'i}^{a'a,\Gamma} + C_{i'i}^{a'a,\Gamma}(-1)^{(i'<i)+(a'>a)} \right) & \Gamma = \Gamma', a \neq a' \text{ and } i \neq i' \\ 2(-1)^{i+n_\Gamma}(-1)^{i'+n_{\Gamma'}} \left(\mathcal{A}_{i'i}^{a'a,\Gamma\Gamma'} + \mathcal{B}_{i'i}^{a'a,\Gamma\Gamma'} \right) & \Gamma \neq \Gamma' \end{cases} \tag{1.16}$$

For the notation, see Section 1.4.3. The Hessian can be more compactly expressed as:

$$\frac{\partial^2 f(\hat{K}=0)}{\partial K_i^{a,\Gamma} \partial K_{i'}^{a'\Gamma'}} = \begin{cases} -2C_0 & \text{if } \Gamma = \Gamma', a = a', i = i' \\ 2(-1)^{i+n_\Gamma}(-1)^{i'+n_{\Gamma'}}\left(\mathcal{D}_{i'i}^{a'a,\Gamma\Gamma'} + \delta_{\Gamma\Gamma'}(1-\delta_{a'a})(1-\delta_{i'i})C_{i'i}^{a'a,\Gamma}(-1)^{(i'<i)+(a'>a)}\right) & \text{otherwise} \end{cases}$$

$$(1.17)$$

1.3.3   Transformation of the wave function

In the optimisation process, $|\Phi\rangle$ varies and we would need the Jacobian and the Hessian at $\hat{K} \neq 0$. The expressions are much more complicated and we avoid this by making a full transformation of $|\Psi_{\text{FCI}}\rangle$ to the new orbital basis. Let $U = e^{-\hat{K}}$ be the matrix that transform the orbital basis:

$$\phi_p = \sum_r \phi'_r U_{rp} \,. \tag{1.18}$$

Given the coefficients C_I of the expansion in the first basis, we want to know the coefficients C'_I such that

$$|\Psi_{\text{FCI}}\rangle = \sum_I C_I |\Phi_I\rangle = \sum_I C'_I |\Phi'_I\rangle \,. \tag{1.19}$$

These are given by:

$$C'_I = \sum_J C_J \det(U_{IJ}) \,, \tag{1.20}$$

where U_{IJ} is the minor of the matrix U with the entries in the rows and columns given by the multi-indices I and J.

Such transformation is the most time consuming step.

Spin and symmetry adapted

In the case that the transformation matrix U is spin and symmetry blocks, the determinant can be calculated as the product of smaller determinants, simplifying Equation (1.20). That is, suppose that the matrix U is given as the direct sum of smaller square matrices (that are the transformation matrix in each spirrep block):

$$\begin{aligned} U &= U^{1\alpha} \oplus U^{2\alpha} \oplus \cdots \oplus U^{g\beta} \\ &= \bigoplus_{\Gamma\sigma} U^{\Gamma\sigma} \,, \end{aligned} \tag{1.21}$$

Then

$$C'_I = \sum_{\substack{J \\ \text{occ}(J)=\text{occ}(I)}} C_J \prod_{\Gamma\sigma} \det(U_{I^{\Gamma\sigma}J^{\Gamma\sigma}}) \,, \tag{1.22}$$

where the sum is only over the multiindices that have the same occupation pattern per spirrep block of I (otherwise $det(U_{IJ}) = 0$, and some of $U_{I^{\Gamma\sigma}J^{\Gamma\sigma}}$ are non square).

1.3.4 Newton-Raphson step

Starting from an orbital basis $\{\phi_p\}$ such that the first determinant (that is, with the first n_α α orbitals and first n_β β occupied) is $|\Phi_0\rangle$, we calculate the Jacobian $\mathbf{J}$ and the Hessian $\mathbf{H}$ as shown above. The Newton step (in the space of the K_i^a parameters) is

$$\mathbf{z} = -\mathbf{H}^{-1}\mathbf{J}\,. \tag{1.23}$$

From this vector, the operator $\hat{K}$ is constructed and the orbital transformation matrix is given by

$$U = e^{-\hat{K}}\,. \tag{1.24}$$

This is done for the α and the β orbitals and the wave function $|\Psi_{\text{FCI}}\rangle$ is transformed to the new orbital basis, by equation (1.20). This proceeds until convergence.

1.4   The Newton-Grassmann method of Absil

This method is based on the work of Absil *et al.*, where they propose a Newton method adapted to the geometry of the Grassmannian. The algorithm is actually on the Stiefel manifold, that is, the manifold of full-rank $K \times n$ matrices. These matrices contain the coefficients of the orbitals in a fixed basis, and thus form a redundant parametrisation for the Grassmannian, in the sense that there are several matrices to represent the same point in Gr, that can be transformed among themselves with multiplication by invertible $n \times n$ matrices (elements of the general linear group of degree n, GL_n).

This algorithm has been submitted to the CNMAC 2020, in the category of "complete works". [ADD REFERENCE IF ACCEPTED]

1.4.1   General formulation from Absil

Thus, given $|\Phi_0\rangle \in Gr$, let Y be the matrix whose entries are the coefficients of $|\Phi_0\rangle$ in some basis of $\mathcal{V}$. This is an element of the Stiefel manifild $ST(n, K)$. Let $f : Gr \to \Re$ be any function we might want to optimise. This is defined on the Grassmannian. We will recall some definitions from Absil.[] The version of this function on the Stiefel manifold is defined as:

$$f_\diamond : ST \quad \to \quad \Re \tag{1.25}$$
$$Y \quad \mapsto \quad f(\text{span}(Y)), \tag{1.26}$$

where $\text{span}(Y)$ is the vector space spanned by the columns of Y (and thus in Gr). The gradient of $f_\diamond$ at Y is the matrix $K \times n$ whose entries are given by:

$$(\text{grad} f_\diamond(Y))_q^p = \frac{\partial f_\diamond(Y)}{\partial Y_q^p}(Y), \tag{1.27}$$

and the directional derivative of a function any smooth F, at x in the direction of y:

$$DF(x)[y] = \frac{d}{dt} F(x + ty)\big|_{t=0}. \tag{1.28}$$

Finally, let

$$\Pi_{W\perp} = I - W(W^T W)^{-1} W^T, \tag{1.29}$$

be the projection into the orthogonal component of W.

The calculation is done now in the Stiefel manifold. According to Absil:

- First solve the following equation for the unknown $\eta_{\diamond Y} \in H_Y = \{Y_\perp K : K \in \Re^{(K-n)\times n}\}$:

$$\Pi_{Y\perp} D \left(\Pi_{.\perp} \text{grad} f_\diamond (\cdot)\right) (Y) [\eta_{\diamond Y}] = -\Pi_{Y\perp} \text{grad} f_\diamond (Y) ; \tag{1.30}$$

- Update $|\Phi_0\rangle \to |\Phi\rangle$ by moving along the geodesic on the Grassmannian in the direction of $\text{span}(\eta_{\diamond Y})$, what is done computing a singular value decomposition of $\eta_{\diamond Y} = U\Sigma V^T$ and getting:

$$|\Phi\rangle = \text{span}(YV \cos\Sigma + U \sin\Sigma). \tag{1.31}$$

1.4.2 **Algorithm 2** Direct application to obtain $|\Phi_{\mathbf{minD}}\rangle$

We will now consider our specific function. Similarly to what we did in the first algorithm, we will just work without getting the absolute value. Thus, we will look for stationary points of:

$$f : Gr \;\rightarrow\; \Re \tag{1.32}$$

$$|\Phi\rangle \;\mapsto\; \langle\Psi_{\text{ext}}|\Phi\rangle \; . \tag{1.33}$$

We will assume that the external wave function, $|\Psi_{\text{ext}}\rangle$, is eigenfunction of S_z, with eigenvalue M_s, and belongs to an irreducible representation (irrep, Γ) of the point group of the molecule. This means that, for all determinants of the expansion in Equation (1.1), the following equations hold:

$$n_\alpha - n_\beta = 2Ms \tag{1.34}$$

$$\bigotimes_q \Gamma(\phi_{I_q}) = \Gamma(|\Psi_{\text{ext}}\rangle) = \Gamma(|\Phi\rangle) \tag{1.35}$$

That is, all determinants have the same Ms (that is half of the difference between alpha and beta orbitals) and belong to the same irrep Γ (that is the direct product of the irrep of all occupied orbitals). We are thus assuming that we are using a basis for $\mathcal{V}$ for which all elements belong to an irrep. Let us further assume that the orbitals are ordered with all alpha coming first, followed by all beta orbitals, and they are ordered by irrep in each of these parts. Under these conditions, the $f_\Diamond$ is calculated as follows:

$$
\begin{aligned}
f_\Diamond(Y) \;&=\; \frac{1}{\sqrt{\det\left(Y^T Y\right)}} \sum_I C_I \det\left(Y\big|_I\right) &(1.36)\\[2ex]
&=\; \frac{1}{\sqrt{\det\left((Y_\alpha)^T Y_\alpha\right)\det\left((Y_\beta)^T Y_\beta\right)}} \sum_I C_I \det\left(Y_\alpha\big|_{I_\alpha}\right)\det\left(Y_\beta\big|_{I_\beta}\right) &(1.37)\\[2ex]
&=\; \frac{1}{\sqrt{\prod_\Gamma \det\left((Y_\alpha^\Gamma)^T Y_\alpha^\Gamma\right)\det\left((Y_\beta^\Gamma)^T Y_\beta^\Gamma\right)}} \sum_{\substack{I\\ \text{occ}(I)=\text{occ}(Y)}} C_I \prod_\Gamma \det\left(Y_\alpha^\Gamma\big|_{I_\alpha^\Gamma}\right)\det\left(Y_\beta^\Gamma\big|_{I_\beta^\Gamma}\right) &(1.38)
\end{aligned}
$$

The above expressions seem very complicated, so let us look at then carefully.

First of all, note that the first expression is general, in the second we consider the separation between α and β orbitals, and in the third we consider the separation between orbitals of different irreps. In all versions, the term with the square root is to guarantee the normalization of the Slater determinant associated to Y. If the coefficients in Y are for an orthonormal set of orbitals, the term in the square root is one. But we note that $f_\Diamond(Y) = f(\text{span}(Y))$ must hold for all Y in the Stiefel manifold, and the normalisation factor is important to use equation (1.30), even if Y is used normalised during computation.

The notation $Y\big|_I$ means the submatrix of Y formed with the rows that are in the multiindex I. For

example, if

$$Y = \begin{pmatrix} Y_1^1 & Y_2^1 & Y_3^1 \\ Y_1^2 & Y_2^2 & Y_3^2 \\ Y_1^3 & Y_2^3 & Y_3^3 \\ Y_1^4 & Y_2^4 & Y_3^4 \\ Y_1^5 & Y_2^5 & Y_3^5 \end{pmatrix} , \tag{1.39}$$

and $I = \{1, 2, 3\}$, then

$$Y\big|_I = \begin{pmatrix} Y_1^1 & Y_2^1 & Y_3^1 \\ Y_1^2 & Y_2^2 & Y_3^2 \\ Y_1^3 & Y_2^3 & Y_3^3 \end{pmatrix} , \tag{1.40}$$

but if $I = \{1, 4, 5\}$, then

$$Y\big|_I = \begin{pmatrix} Y_1^1 & Y_2^1 & Y_3^1 \\ Y_1^4 & Y_2^4 & Y_3^4 \\ Y_1^5 & Y_2^5 & Y_3^5 \end{pmatrix} , \tag{1.41}$$

and so on. Thus, the first formula says that, after taking the normalisation into account, $f_\Diamond$ is calculated by summing the contribution of all determinants that contribute to $|\Psi_{\text{ext}}\rangle$, and this contribution is the CI coefficient times the determinant of $Y\big|_I$.

Now, if the orbitals of a Slater determinant $|\Phi\rangle$ can be divided into α and β orbitals, that do not mix, the matrix Y is divided in two blocks, one for each spin:

$$Y = \begin{pmatrix} Y_\alpha & 0 \\ 0 & Y_\beta \end{pmatrix} = Y_\alpha \oplus Y_\beta . \tag{1.42}$$

The second equation follows easily from the first when we use this division into blocks, because the determinant of block diagonal matrices is the product of the determinant of each block. Furthermore, when we consider symmetry the matrix is further divided:

$$Y = \begin{pmatrix} Y_\alpha^{\Gamma_1} & 0 & \dots & 0 \\ 0 & Y_\alpha^{\Gamma_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & Y_\beta^{\Gamma_g} \end{pmatrix} = Y_\alpha^{\Gamma_1} \oplus Y_\alpha^{\Gamma_2} \oplus \cdots \oplus Y_\beta^{\Gamma_g} , \tag{1.43}$$

leading to the last formula. In this case, however, the matrices Y_σ^Γ might not be square, because by fixing the irrep of the wave function (and thus of its determinants) does not fix the occupation in each irrep block: it is possible to fulfil equation (1.35) with more than one pattern in the occupation of the irreps Γ. In the summation over I, if this multi-index is associated to a determinant with an occupation per irrep different than the occupation if Y, some of the blocks Y_σ^Γ are not square, although the direct sum of all of them is again square (because the total number of electrons does not change). However, one can show that the determinant of such block matrices are zero, and thus they do not contribute to $f_\Diamond$. This is indicated in the summation by "$\text{occ}(I) = \text{occ}(Y)$", that is, we can consider only the terms with the same occupation per irrep of Y.

The notation I_σ^Γ indicates the part of the multi-index I associated to σ and Γ. Because it is obvious that we can take submatrices of Y_α^Γ only restricting with subindices with same spin and irrep, the notation $Y_\alpha^\Gamma\big|_{I_\alpha^\Gamma}$ is rather redundant. Hence we will use $Y_\alpha^\Gamma\big|_I$ or $Y\big|_{I_\alpha^\Gamma}$ only.

Because of the separation in spin and irrep blocks, we can show that the unknown matrix $\eta_{\Diamond Y}$ has the same division as the matrix Y:

$$\eta_{\Diamond Y} = \begin{pmatrix} \eta_\alpha^{\Gamma_1} & 0 & \dots & 0 \\ 0 & \eta_\alpha^{\Gamma_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \eta_\beta^{\Gamma_g} \end{pmatrix} = \eta_\alpha^{\Gamma_1} \oplus \eta_\alpha^{\Gamma_2} \oplus \dots \oplus \eta_\beta^{\Gamma_g}, \tag{1.44}$$

where we dropped the subscript $\Diamond Y$ in the blocks for a cleaner notation.

After a big exercise of linear algebra, we obtain the following system of linear equations as the main equation of this method (i.e. Equation (1.30) using the present function f):

$$\left(\mathbf{X}_{\sigma\Gamma}^{\sigma'\Gamma'}\right)_{qs}^{pr}\left(\eta_{\sigma'\Gamma'}\right)_s^r = \left(\mathbf{C}_{\sigma\Gamma}\right)_q^p, \tag{1.45}$$

with

$$\left(\mathbf{C}_{\sigma\Gamma}\right)_q^p = \sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\{\sigma',\Gamma'\}\neq\{\sigma,\Gamma\}} F_{I_{\sigma'}^{\Gamma'}}\right) \left(G_{I_\sigma^\Gamma} - F_{I_\sigma^\Gamma}Y_\sigma^\Gamma\right)_q^p \tag{1.46}$$

$$= \left(\Pi_{Y_\sigma^\Gamma\perp}\right)_{\bar p}^p \sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\{\sigma',\Gamma'\}\neq\{\sigma,\Gamma\}} F_{I_{\sigma'}^{\Gamma'}}\right) \left(G_{I_\sigma^\Gamma}\right)_q^{\bar p} \tag{1.47}$$

$$\left(\mathbf{X}_{\sigma\Gamma}^{\sigma\Gamma}\right)_{qs}^{pr} = \sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\{\sigma',\Gamma'\}\neq\{\sigma,\Gamma\}} F_{I_{\sigma'}^{\Gamma'}}\right) \left\{\left(G_{I_\sigma^\Gamma} - F_{I_\sigma^\Gamma}Y_\sigma^\Gamma\right)_q^p \left(Y_\sigma^\Gamma\right)_s^r - \left(\Pi_{Y_\sigma^\Gamma\perp}\right)_{\bar p}^p \left(\tilde H_{I_\sigma^\Gamma}\right)_{qs}^{\bar p r}\right\} \tag{1.48}$$

$$= \left(\Pi_{Y_\sigma^\Gamma\perp}\right)_{\bar p}^p \sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\{\sigma',\Gamma'\}\neq\{\sigma,\Gamma\}} F_{I_{\sigma'}^{\Gamma'}}\right) \left\{\left(G_{I_\sigma^\Gamma}\right)_q^{\bar p}\left(Y_\sigma^\Gamma\right)_s^r - \left(\tilde H_{I_\sigma^\Gamma}\right)_{qs}^{\bar p r}\right\} \tag{1.49}$$

$$\left(\mathbf{X}_{\sigma\Gamma}^{\sigma'\Gamma'}\right)_{qs}^{pr} = -\sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\substack{\{\sigma'',\Gamma''\}\neq\{\sigma',\Gamma'\} \\ \{\sigma'',\Gamma''\}\neq\{\sigma,\Gamma\}}} F_{I_{\sigma''}^{\Gamma''}}\right) \left(G_{I_\sigma^\Gamma} - F_{I_\sigma^\Gamma}Y_\sigma^\Gamma\right)_q^p \left(G_{I_{\sigma'}^{\Gamma'}} - F_{I_{\sigma'}^{\Gamma'}}Y_{\sigma'}^{\Gamma'}\right)_s^r \tag{1.50}$$

$$= -\left(\Pi_{Y_\sigma^\Gamma\perp}\right)_{\bar p}^p\left(\Pi_{Y_{\sigma'}^{\Gamma'}\perp}\right)_{\bar r}^r \sum_{\substack{I \\ \text{occ}(I)=\text{occ}(Y)}} C_I \left(\prod_{\substack{\{\sigma'',\Gamma''\}\neq\{\sigma',\Gamma'\} \\ \{\sigma'',\Gamma''\}\neq\{\sigma,\Gamma\}}} F_{I_{\sigma''}^{\Gamma''}}\right) \left(G_{I_\sigma^\Gamma}\right)_q^{\bar p}\left(G_{I_{\sigma'}^{\Gamma'}}\right)_s^{\bar r} \tag{1.51}$$

where the last equations hold for $\{\sigma,\Gamma\} \neq \{\sigma',\Gamma'\}$ and:

$$F_{I_\sigma^\Gamma} = \det\left(\left.Y\right|_{I_\sigma^\Gamma}\right) \tag{1.52}$$

$$\left(G_{I_\sigma^\Gamma}\right)_q^p = \det\left(Y_\sigma^\Gamma \overset{q}{\leftarrow} e_p\right)\Big|_I \tag{1.53}$$

$$\left(H_{I_\sigma^\Gamma}\right)_{qs}^{pr} = \det\left(Y_\sigma^\Gamma \overset{q}{\leftarrow} e_p \overset{s}{\leftarrow} e_r\right)\Big|_I \tag{1.54}$$

$$\left(\tilde{H}_{I_\sigma^\Gamma}\right)_{qs}^{pr} = \begin{cases} \left(H_{I_\sigma^\Gamma}\right)_{qs}^{pr} & \text{if } s \neq q \\ -F_{I_\sigma^\Gamma}\delta_{pr} & \text{otherwise} \end{cases} \tag{1.55}$$

The notation $A \overset{q}{\leftarrow} b$ means the matrix A with the q-th column replaced by the vector b. The vector e_p is the p-th element of the canonical basis (0 in every entry, except for the p-th entry where it is 1).

1.4.3   **Algorithm 3** The case of a CISD wave function

Although Algorithm 2 already offers an improved procedure to calculate $|\Psi_{\text{minD}}\rangle$, as compared to Algorithm 1, it is still too general. It can be made much more efficient if the equations are adapted to a specific kind of wave function. In this section we will describe the equations for $|\Psi_{\text{ext}}\rangle$ being of the type single reference, closed shell, restricted configuration interaction with single and double excitations: a CISD wave function.

The structure of a CISD wave function

Let $|\Phi_0\rangle$ be the reference determinant of closed shell and restricted type: It has even number of electrons, with the same occupied α and β orbitals:

$$\begin{aligned}
|\Phi_0\rangle &= \phi_1^{\Gamma=1} \wedge \phi_2^{\Gamma=1} \wedge \ldots \wedge \phi_{n_g}^{\Gamma=g} \wedge \overline{\phi}_1^{\Gamma=1} \wedge \overline{\phi}_2^{\Gamma=1} \wedge \ldots \wedge \overline{\phi}_{n_g}^{\Gamma=g} \tag{1.56}\\
&= |\Phi_0\rangle_1 \wedge |\Phi_0\rangle_2 \wedge \ldots \wedge |\Phi_0\rangle_g \wedge \left|\overline{\Phi}_0\right\rangle_1 \wedge \left|\overline{\Phi}_0\right\rangle_2 \wedge \ldots \wedge \left|\overline{\Phi}_0\right\rangle_g \tag{1.57}\\
&= |\Phi_0\rangle \wedge \left|\overline{\Phi}_0\right\rangle . \tag{1.58}
\end{aligned}$$

As can be seen in equations above, we will use the convention of an over line to indicate that it corresponds to β spin, whereas no over line indicates α spin. Furthermore, $|\Phi_0\rangle_\Gamma$ corresponds to the part of $|\Phi_0\rangle$ for irrep Γ and α spin (that is, the subspace of $\mathcal{V}$ spanned by the occupied alpha orbitals of irrep Γ in the reference). $\left|\overline{\Phi}_0\right\rangle_\Gamma$ is the same for the β part. We will also use the standard convention of denoting excited Slater determinants by subscripts for holes ("excited from") and superscripts for particles ("excited to"). Thus, for example:

$$|\Phi_i^a\rangle_\Gamma = \phi_1^\Gamma \wedge \ldots \wedge \hat{\phi}_i^\Gamma \wedge \ldots \wedge \phi_{n_\Gamma}^\Gamma \wedge \phi_a^\Gamma \tag{1.59}$$

$$\left|\Phi_{ij}^{ab}\right\rangle_\Gamma = \phi_1^\Gamma \wedge \ldots \wedge \hat{\phi}_j^\Gamma \wedge \ldots \wedge \hat{\phi}_i^\Gamma \wedge \ldots \wedge \phi_{n_\Gamma}^\Gamma \wedge \phi_b^\Gamma \wedge \phi_a^\Gamma \tag{1.60}$$

$$|\Phi_i\rangle_\Gamma = \phi_1^\Gamma \wedge \ldots \wedge \hat{\phi}_i^\Gamma \wedge \ldots \wedge \phi_{n_\Gamma}^\Gamma \tag{1.61}$$

$$\left|\Phi^{ab}\right\rangle_\Gamma = \phi_1^\Gamma \wedge \ldots \ldots \wedge \phi_{n_\Gamma}^\Gamma \wedge \phi_b^\Gamma \wedge \phi_a^\Gamma. \tag{1.62}$$

In these equations we are using the convention that a hat (ˆ) over an orbital indicates its absence in the product. Note that in the last two cases the number of electrons in $\left|\Phi\right\rangle_\Gamma$ is not preserved: These cases appear in the wave function, as discussed in Sect. 1.4.2.

The restricted closed shell CISD wave function is given as:

$$
\begin{aligned}
|\Psi_{\mathrm{CISD}}\rangle =& C_0 |\Phi_0\rangle \\
&+ \sum_\Gamma \sum_{(i,a)\in\Gamma} C_i^{a,\Gamma} \Big(|\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_i^a\rangle_\Gamma \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle \\
&\qquad\qquad\qquad + |\Phi_0\rangle \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge |\overline{\Phi}_i^a\rangle_\Gamma \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \Big) \\
&+ \sum_\Gamma \sum_{\substack{(i>j)\\(a>b)}\}\in\Gamma} C_{ij}^{ab,\Gamma} \Big(|\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_{ij}^{ab}\rangle_\Gamma \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle \\
&\qquad\qquad\qquad + |\Phi_0\rangle \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge \left|\overline{\Phi}_{ij}^{ab}\right\rangle_\Gamma \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \Big) \\
&+ \sum_\Gamma \sum_{\substack{(i,a)\\(j,b)}\}\in\Gamma} \mathcal{C}_{ij}^{ab,\Gamma} |\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_i^a\rangle_\Gamma \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge \left|\overline{\Phi}_j^b\right\rangle_\Gamma \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma\\(j,b)\in\Gamma'}} \mathcal{A}_{ij}^{ab,\Gamma\Gamma'} \Big(|\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_j^b\rangle_{\Gamma'} \wedge \ldots \wedge |\Phi_i^a\rangle_\Gamma \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle \\
&\qquad\qquad\qquad + |\Phi_0\rangle \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge \left|\overline{\Phi}_j^b\right\rangle_{\Gamma'} \wedge \ldots \wedge |\overline{\Phi}_a^a\rangle_\Gamma \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \Big) \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma\\(j,b)\in\Gamma'}} \mathcal{B}_{ij}^{ab,\Gamma\Gamma'} \Big(|\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_j^b\rangle_{\Gamma'} \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge |\overline{\Phi}_i^a\rangle_\Gamma \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \\
&\qquad\qquad\qquad + |\Phi_0\rangle_1 \wedge \ldots \wedge |\Phi_i^a\rangle_\Gamma \wedge \ldots \wedge |\Phi_0\rangle_g \wedge |\overline{\Phi}_0\rangle_1 \wedge \ldots \wedge |\overline{\Phi}_a^a\rangle_{\Gamma'} \wedge \ldots \wedge |\overline{\Phi}_0\rangle_g \Big) \\
&\quad \sum_{\substack{I \text{ doubles over } |\Phi_0\rangle\\ \mathrm{occ}(I)\neq\mathrm{occ}(0)}} C_I |\Phi_I\rangle \\
=& C_0 |\Phi_0\rangle \\
&+ \sum_\Gamma \sum_{(i,a)\in\Gamma} C_i^{a,\Gamma} \Big(|\Phi_i^a\rangle_\Gamma + \left|\overline{\Phi}_i^a\right\rangle_\Gamma \Big) \\
&+ \sum_\Gamma \sum_{\substack{(i>j)\\(a>b)}\}\in\Gamma} C_{ij}^{ab,\Gamma} \Big(|\Phi_{ij}^{ab}\rangle_\Gamma + \left|\overline{\Phi}_{ij}^{ab}\right\rangle_\Gamma \Big) \\
&+ \sum_\Gamma \sum_{\substack{(i,a)\\(j,b)}\}\in\Gamma} \mathcal{C}_{ij}^{ab,\Gamma} |\Phi_i^a\rangle_\Gamma \ldots \left|\overline{\Phi}_j^b\right\rangle_\Gamma \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma\\(j,b)\in\Gamma'}} \mathcal{A}_{ij}^{ab,\Gamma\Gamma'} \Big(|\Phi_j^b\rangle_{\Gamma'} \ldots |\Phi_i^a\rangle_\Gamma + \left|\overline{\Phi}_j^b\right\rangle_{\Gamma'} \ldots |\overline{\Phi}_i^a\rangle_\Gamma \Big) \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma\\(j,b)\in\Gamma'}} \mathcal{B}_{ij}^{ab,\Gamma\Gamma'} \Big(|\Phi_j^b\rangle_{\Gamma'} \ldots |\overline{\Phi}_i^a\rangle_\Gamma + |\Phi_i^a\rangle_\Gamma \ldots \left|\overline{\Phi}_j^b\right\rangle_{\Gamma'} \Big) \\
&\quad \sum_{\substack{I \text{ doubles over } |\Phi_0\rangle\\ \mathrm{occ}(I)\neq\mathrm{occ}(0)}} C_I |\Phi_I\rangle \, .
\end{aligned}
$$

$$(1.63)$$

In these equations, the first term is the contribution of reference wave function. The second is the

contribution of single excitations, with corresponding α and β excitations having the same coefficient, as it is a restricted wave function over restricted reference. Third term is the contribution of double excitations that are totally within the same *spirrep* block (spin and irrep). For these, we obviously cannot have $i = j$ or $a = b$, and the restrictions in the summation guarantee that we count each excited determinant once. Fourth term considers the contribution of double excitations within the same irrep, but made of a single excitation in α and a single excitation in β. There is no restriction in the summation in this case, but because $|\Psi_{\mathrm{CISD}}\rangle$ is a restricted wave function, the coefficients obey $\mathcal{C}_{ij}^{ab,\Gamma} = \mathcal{C}_{ji}^{ba,\Gamma}$. Fifth and sixth terms are the contributions of double excitations that are made of single excitations in different irreps: The fifth term is for these both single excitations happening in the same spin, whereas in the sixth term these single excitations are in opposite spins. These are all the terms that have, for all spirreps, the same number of electrons as in reference wave function. Determinants that have some irrep with number of electrons different than in the reference determinant are collected in the last term, and they do not contribute neither to $f_\diamond$ nor to the matrices used in the optimisation. In the second of the above equations we have compacted the notation, showing only the spirrep block(s) that differ from the corresponding reference block.

The equations

Using the wave function of Equation (1.63) in the equations of Algorithm 2, a number of simplifications and collection of similar terms can be made, with final equations described in this section. We first start by defining some intermediates:

$$\mathcal{F}_0 = \prod_\Gamma F_{I_0^\Gamma}^2 \tag{1.64}$$

$$\mathcal{F}_0^\Gamma = \prod_{\Gamma' \neq \Gamma} F_{I_0^{\Gamma'}}^2 \tag{1.65}$$

$$\mathcal{F}_0^{\Gamma\Gamma'} = \prod_{\substack{\Gamma'' \neq \Gamma \\ \Gamma'' \neq \Gamma'}} F_{I_0^{\Gamma''}}^2 , \tag{1.66}$$

with analogous definitions for $\mathcal{F}_0^{\Gamma\Gamma'\Gamma''}$ and so on.

$$\mathcal{D}_{ij}^{ab,\Gamma\Gamma} = \mathcal{C}_{ij}^{ab,\Gamma} \tag{1.67}$$

$$\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} = \mathcal{A}_{ij}^{ab,\Gamma\Gamma'} + \mathcal{B}_{ij}^{ab,\Gamma\Gamma'} \quad \text{for } \Gamma \neq \Gamma', \tag{1.68}$$

that is, the coefficients of double excitations coming from single excitations at different spirrep blocks are merged in a single quantity $\mathcal{D}$. As it will be seen below, $\mathcal{A}_{ij}^{ab,\Gamma\Gamma'}$ and $\mathcal{B}_{ij}^{ab,\Gamma\Gamma'}$ always appear summed.

$$\mathcal{K}^{\Gamma\Gamma'} = 2 \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'} F_{I_0^\Gamma} F_{I_0^{\Gamma'}} F_{I_i^{a,\Gamma}} F_{I_j^{b,\Gamma'}} \quad \text{for } \Gamma \neq \Gamma' \tag{1.69}$$

Here, I_0^Γ is the multi-index of the reference for irrep Γ, whereas $I_i^{a,\Gamma}$ is the multi-index for the single excitation from i to a, also in Γ. The other type of multi-index that will appear is, obviously, $I_{ij}^{ab,\Gamma}$, for double excitations.

$$\mathcal{L}^\Gamma = 2F_{I_0^\Gamma}\left(\sum_{(i,a)\in\Gamma} C_i^{a,\Gamma} F_{I_i^{a,\Gamma}} + \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma} F_{I_{ij}^{ab,\Gamma}}\right) + \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}} F_{I_j^{b,\Gamma}} \tag{1.70}$$

Note the factor 2: it is there because we have the same contribution for α and β excitations totally within a spirerp block. With the so far defined quantities, we are able to calculate $f_\diamond(Y)$, assuming Y orthonormal:

$$f_\diamond(Y) = C_0 \mathcal{F}_0 + \sum_\Gamma \mathcal{F}_0^\Gamma \mathcal{L}^\Gamma + \sum_{\Gamma>\Gamma'} \mathcal{F}_0^{\Gamma\Gamma'} \mathcal{K}^{\Gamma\Gamma'}. \tag{1.71}$$

It is not difficult to understand the origin of each term of Equation (1.71) (compare to Equation (1.36)): The first is clearly the contribution from the reference determinant, that is the product of all $F_{I_0^\Gamma}$, for all Γ and for each spin. Since it is a closed shell restricted wave function, this is just $\mathcal{F}_0$. The second term of Equation (1.71) is the contribution of all excitations within the same irrep: for all such excitations, the other irrep blocks contribute with a $F_{I_0^{\Gamma'}}^2$, that form a common $\mathcal{F}_0^\Gamma$; the contribution of the irrep in question is the coefficient, times an appropriate F_{I^Γ}, as can be seen in Equation (1.70). Finally, last term is the contribution of excitations at mixed irreps (say Γ and Γ'): each determinant contribute with one $F_{I_0^\Gamma}$, one $F_{I_0^{\Gamma'}}$ (from the spins where no excitations occurred, whichever they are), and the F_{I^Γ} and $F_{I^{\Gamma'}}$ of corresponding single excitations. This is clearly seen in (1.69), and the contribution of remaining irreps forms $\mathcal{F}_0^{\Gamma\Gamma'}$.

For the matrices $\mathbf{X}$ and $\mathbf{C}$, the interpretation is more elaborated but follows similar line. We will define more quantities. These quantities are matrices with shape (K_Γ, n^Γ), such as Y^Γ and $\mathbf{C}_{I^\Gamma}$, or tensors [1] of shape $(K_\Gamma, n^\Gamma, K_{\Gamma'}, n^{\Gamma'})$, such as $H_{I^{\Gamma\Gamma'}}$ and $\mathbf{X}_{\Gamma'}^\Gamma$. For completeness and for reference, we will give the expressions for the multi-dimensional quantities and for the individual elements.

$$\hat{\mathbf{G}}_{ia}^\Gamma = F_{I_0^\Gamma}\mathbf{G}_{I_i^{a,\Gamma}} + F_{I_i^{a,\Gamma}}\mathbf{G}_{I_0^\Gamma} \tag{1.72}$$

$$\left(\hat{\mathbf{G}}_{ia}^\Gamma\right)_q^p = F_{I_0^\Gamma}\left(\mathbf{G}_{I_i^{a,\Gamma}}\right)_q^p + F_{I_i^{a,\Gamma}}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^p \tag{1.73}$$

$$\overline{C}_i^{a,\Gamma} = C_i^{a,\Gamma} + \sum_{\overline{\Gamma}\neq\Gamma}\frac{1}{F_{I_0^{\overline{\Gamma}}}}\sum_{(j,b)\in\overline{\Gamma}} F_{I_j^{b,\overline{\Gamma}}}\mathcal{D}_{ij}^{ab,\Gamma\overline{\Gamma}} \tag{1.74}$$

$$\begin{aligned}
\boldsymbol{\mathcal{M}}^\Gamma =\,& C_0 F_{I_0^\Gamma}\mathbf{G}_{I_0^\Gamma} \\
&+ \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\hat{\mathbf{G}}_{ia}^\Gamma \\
&+ \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma}\left(F_{I_0^\Gamma}\mathbf{G}_{I_{ij}^{ab,\Gamma}} + F_{I_{ij}^{ab,\Gamma}}\mathbf{G}_{I_0^\Gamma}\right) \\
&+ \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}}\mathbf{G}_{I_j^{b,\Gamma}}
\end{aligned} \tag{1.75}$$

---

[1] Tensors in the sense of multidimensional array objects, such as a numpy array. Probably the indices convention used here does not follow the rules of contravariant and covariant tensors, although I tried to use a meaningful convention.

$$
\begin{aligned}
\left(\boldsymbol{\mathcal{M}}^\Gamma\right)_q^p =\;& C_0 F_{I_0^\Gamma}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^p \\
& + \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\left(\hat{\mathbf{G}}_{ia}^\Gamma\right)_q^p \\
& + \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma}\left(F_{I_0^\Gamma}\left(\mathbf{G}_{I_{ij}^{ab,\Gamma}}\right)_q^p + F_{I_{ij}^{ab,\Gamma}}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^p\right) \\
& + \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}}\left(\mathbf{G}_{I_j^{b,\Gamma}}\right)_q^p
\end{aligned}
\tag{1.76}
$$

$$
\begin{aligned}
\mathscr{G}^\Gamma =\;& C_0 \mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_0^\Gamma} \\
& + \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\left(\mathbf{G}_{I_i^{a,\Gamma}} \otimes \mathbf{G}_{I_0^\Gamma} + \mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_i^{a,\Gamma}}\right) \\
& + \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma}\left(\mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_{ij}^{ab,\Gamma}} + \mathbf{G}_{I_{ij}^{ab,\Gamma}} \otimes \mathbf{G}_{I_0^\Gamma}\right) \\
& + \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} \mathbf{G}_{I_i^{a,\Gamma}} \otimes \mathbf{G}_{I_j^{b,\Gamma}}
\end{aligned}
\tag{1.77}
$$

$$
\begin{aligned}
\left(\mathscr{G}^\Gamma\right)_{qs}^{pr} =\;& C_0\left(\mathbf{G}_{I_0^\Gamma}\right)_q^p\left(\mathbf{G}_{I_0^\Gamma}\right)_s^r \\
& + \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\left(\left(\mathbf{G}_{I_i^{a,\Gamma}}\right)_q^p\left(\mathbf{G}_{I_0^\Gamma}\right)_s^r + \left(\mathbf{G}_{I_0^\Gamma}\right)_q^p\left(\mathbf{G}_{I_i^{a,\Gamma}}\right)_s^r\right) \\
& + \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma}\left(\left(\mathbf{G}_{I_0^\Gamma}\right)_q^p\left(\mathbf{G}_{I_{ij}^{ab,\Gamma}}\right)_s^r + \left(\mathbf{G}_{I_{ij}^{ab,\Gamma}}\right)_q^p\left(\mathbf{G}_{I_0^\Gamma}\right)_s^r\right) \\
& + \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma}\left(\mathbf{G}_{I_i^{a,\Gamma}}\right)_q^p\left(\mathbf{G}_{I_j^{b,\Gamma}}\right)_s^r
\end{aligned}
\tag{1.78}
$$

$$
\begin{aligned}
\mathscr{H}^\Gamma =\;& C_0 F_{I_0^\Gamma} \tilde{H}_{I_0^\Gamma} \\
& + \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\left(F_{I_0^\Gamma}\tilde{H}_{I_i^{a,\Gamma}} + F_{I_i^{a,\Gamma}}\tilde{H}_{I_0^\Gamma}\right) \\
& + \sum_{\substack{(i>j)\\(a>b)}\in\Gamma} C_{ij}^{ab,\Gamma}\left(F_{I_0^\Gamma}\tilde{H}_{I_{ij}^{ab,\Gamma}} + F_{I_{ij}^{ab,\Gamma}}\tilde{H}_{I_0^\Gamma}\right) \\
& + \sum_{\substack{(i,a)\\(j,b)}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}}\tilde{H}_{I_j^{b,\Gamma}}
\end{aligned}
\tag{1.79}
$$

$$
\begin{aligned}
\left(\mathscr{H}^\Gamma\right)_q^p =\; & C_0 F_{I_0^\Gamma}\left(\tilde H_{I_0^\Gamma}\right)_{qs}^{pr} \\
& + \sum_{(i,a)\in\Gamma} \overline{C}_i^{a,\Gamma}\left(F_{I_0^\Gamma}\left(\tilde H_{I_i^{a,\Gamma}}\right)_{qs}^{pr} + F_{I_i^{a,\Gamma}}\left(\tilde H_{I_0^\Gamma}\right)_{qs}^{pr}\right) \\
& + \sum_{\left.\begin{smallmatrix}(i>j)\\(a>b)\end{smallmatrix}\right\}\in\Gamma} C_{ij}^{ab,\Gamma}\left(F_{I_0^\Gamma}\left(\tilde H_{I_{ij}^{ab,\Gamma}}\right)_{qs}^{pr} + F_{I_{ij}^{ab,\Gamma}}\left(\tilde H_{I_0^\Gamma}\right)_{qs}^{pr}\right) \\
& + \sum_{\left.\begin{smallmatrix}(i,a)\\(j,b)\end{smallmatrix}\right\}\in\Gamma} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}}\left(\tilde H_{I_j^{b,\Gamma}}\right)_{qs}^{pr}
\end{aligned}
\tag{1.80}
$$

Although these equations seem very complicated and too big, the terms are very similar to what we have for $\mathcal{L}^\Gamma$: a term for the reference determinant (okay, that is not in $\mathcal{L}^\Gamma$, but in $f_\diamond$), a term for "single excitations", another for double excitations in a spirrep block, and a term for double excitations in Γ, but coming from single excitations in opposite spins. However, the term for "single excitations" contains a contribution from double excitations that occur partially in Γ and partially in another irrep, $\overline{\Gamma}$, that is indeed a single excitation in Γ.

Finally:

$$
\mathbf{C}_\Gamma = \Pi_{Y^\Gamma\perp}\left\{\mathcal{F}_0^\Gamma\mathcal{M}^\Gamma + \left(\sum_{\overline{\Gamma}\neq\Gamma}\mathcal{F}_0^{\Gamma\overline{\Gamma}}\mathcal{L}^{\overline{\Gamma}} + \sum_{\substack{\overline{\overline{\Gamma}}>\overline{\Gamma}\\\overline{\Gamma}\neq\Gamma\\\overline{\overline{\Gamma}}\neq\Gamma}}\mathcal{F}_0^{\Gamma\overline{\Gamma}\overline{\overline{\Gamma}}}\mathcal{K}^{\overline{\Gamma}\overline{\overline{\Gamma}}}\right)F_{I_0^\Gamma}\mathbf{G}_{I_0^\Gamma}\right\}
\tag{1.81}
$$

$$
\left(\mathbf{C}_\Gamma\right)_q^p = \left(\Pi_{Y^\Gamma\perp}\right)_{\bar p}^p\left\{\mathcal{F}_0^\Gamma\left(\mathcal{M}^\Gamma\right)_q^{\bar p} + \left(\sum_{\overline{\Gamma}\neq\Gamma}\mathcal{F}_0^{\Gamma\overline{\Gamma}}\mathcal{L}^{\overline{\Gamma}} + \sum_{\substack{\overline{\overline{\Gamma}}>\overline{\Gamma}\\\overline{\Gamma}\neq\Gamma\\\overline{\overline{\Gamma}}\neq\Gamma}}\mathcal{F}_0^{\Gamma\overline{\Gamma}\overline{\overline{\Gamma}}}\mathcal{K}^{\overline{\Gamma}\overline{\overline{\Gamma}}}\right)F_{I_0^\Gamma}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar p}\right\}
\tag{1.82}
$$

$$
\begin{aligned}
\mathbf{X}_\Gamma^\Gamma =\; & \left(\Pi_{Y^\Gamma\perp}\otimes\mathbb{1}\right)\left\{\mathcal{F}_0^\Gamma\left(\mathcal{M}^\Gamma\otimes Y^\Gamma - \left(\mathbb{1}\otimes\Pi_{Y^\Gamma\perp}\right)\mathscr{G}^\Gamma - \mathscr{H}^\Gamma\right)\right. \\
& \left. + \left(\sum_{\overline{\Gamma}\neq\Gamma}\mathcal{F}_0^{\Gamma\overline{\Gamma}}\mathcal{L}^{\overline{\Gamma}} + \sum_{\substack{\overline{\overline{\Gamma}}>\overline{\Gamma}\\\overline{\Gamma}\neq\Gamma\\\overline{\overline{\Gamma}}\neq\Gamma}}\mathcal{F}_0^{\Gamma\overline{\Gamma}\overline{\overline{\Gamma}}}\mathcal{K}^{\overline{\Gamma}\overline{\overline{\Gamma}}}\right)\left(F_{I_0^\Gamma}\mathbf{G}_{I_0^\Gamma}\otimes Y^\Gamma - \mathbf{G}_{I_0^\Gamma}\otimes\left(\Pi_{Y^\Gamma\perp}\mathbf{G}_{I_0^\Gamma}\right) - F_{I_0^\Gamma}\tilde H_{I_0^\Gamma}\right)\right\}
\end{aligned}
\tag{1.83}
$$

$$
\begin{aligned}
\left(\mathbf{X}_\Gamma^\Gamma\right)_{qs}^{pr} =\; & \left(\Pi_{Y^\Gamma\perp}\right)_{\bar p}^p\left\{\mathcal{F}_0^\Gamma\left(\left(\mathcal{M}^\Gamma\right)_q^{\bar p}\left(Y^\Gamma\right)_s^r - \left(\Pi_{Y^\Gamma\perp}\right)_{\bar r}^r\left(\mathscr{G}^\Gamma\right)_{qs}^{\bar p\bar r} - \left(\mathscr{H}^\Gamma\right)_{qs}^{\bar p r}\right)\right. \\
& \left. + \left(\sum_{\overline{\Gamma}\neq\Gamma}\mathcal{F}_0^{\Gamma\overline{\Gamma}}\mathcal{L}^{\overline{\Gamma}} + \sum_{\substack{\overline{\overline{\Gamma}}>\overline{\Gamma}\\\overline{\Gamma}\neq\Gamma\\\overline{\overline{\Gamma}}\neq\Gamma}}\mathcal{F}_0^{\Gamma\overline{\Gamma}\overline{\overline{\Gamma}}}\mathcal{K}^{\overline{\Gamma}\overline{\overline{\Gamma}}}\right)\left(F_{I_0^\Gamma}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar p}\left(Y^\Gamma\right)_s^r - \left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar p}\left(\Pi_{Y^\Gamma\perp}\right)_{\bar r}^r\left(\mathbf{G}_{I_0^\Gamma}\right)_s^{\bar r} - F_{I_0^\Gamma}\left(\tilde H_{I_0^\Gamma}\right)_{qs}^{\bar p r}\right)\right\}
\end{aligned}
\tag{1.84}
$$

$$
\begin{aligned}
\mathbf{X}_\Gamma^{\Gamma'} = &- 2(\Pi_{Y^\Gamma\perp} \otimes \Pi_{Y^{\Gamma'}\perp})\Bigg\{ \mathcal{F}_0^{\Gamma\Gamma'}\Bigg\{ F_{I_0^{\Gamma'}}\boldsymbol{\mathcal{M}}^\Gamma \otimes \mathbf{G}_{I_0^{\Gamma'}} + F_{I_0^\Gamma}\mathbf{G}_{I_0^\Gamma} \otimes \boldsymbol{\mathcal{M}}^{\Gamma'} \\
&+ \frac{1}{2} \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'}\hat{\mathbf{G}}_{ia}^\Gamma \otimes \hat{\mathbf{G}}_{jb}^{\Gamma'} \\
&- \Bigg(\sum_{(i,a)\in\Gamma} \hat{\mathbf{G}}_{ia}^\Gamma \sum_{(j,b)\in\Gamma'} F_{I_j^{b,\Gamma'}}\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \Bigg) \otimes \mathbf{G}_{I_0^{\Gamma'}} \\
&- \mathbf{G}_{I_0^\Gamma} \otimes \Bigg(\sum_{(j,b)\in\Gamma'} \hat{\mathbf{G}}_{jb}^{\Gamma'} \sum_{(i,a)\in\Gamma} F_{I_i^{a,\Gamma'}}\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \Bigg) \Bigg\} \\
&+ \Bigg(-C_0\mathcal{F}_0^{\Gamma\Gamma'} + \sum_{\bar{\Gamma}\neq\Gamma,\Gamma'} \mathcal{F}_0^{\Gamma\Gamma'\bar{\Gamma}}\mathcal{L}^{\bar{\Gamma}} + \sum_{\substack{\bar{\bar{\Gamma}}>\bar{\Gamma} \\ \bar{\Gamma}\neq\Gamma,\Gamma' \\ \bar{\bar{\Gamma}}\neq\Gamma,\Gamma'}} \mathcal{F}_0^{\Gamma\Gamma'\bar{\Gamma}\bar{\bar{\Gamma}}}\mathcal{K}^{\bar{\Gamma}\bar{\bar{\Gamma}}} \Bigg)F_{I_0^\Gamma}F_{I_0^{\Gamma'}}\mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_0^{\Gamma'}} \Bigg\}
\end{aligned}
\tag{1.85}
$$

$$
\begin{aligned}
\left(\mathbf{X}_\Gamma^{\Gamma'}\right)_{qs}^{pr} = &- 2\left(\Pi_{Y^\Gamma\perp}\right)_{\bar{p}}^p\left(\Pi_{Y^{\Gamma'}\perp}\right)_{\bar{r}}^r\Bigg\{ \mathcal{F}_0^{\Gamma\Gamma'}\Bigg\{ F_{I_0^{\Gamma'}}\left(\boldsymbol{\mathcal{M}}^\Gamma\right)_q^{\bar{p}}\left(\mathbf{G}_{I_0^{\Gamma'}}\right)_s^{\bar{r}} + F_{I_0^\Gamma}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar{p}}\left(\boldsymbol{\mathcal{M}}^{\Gamma'}\right)_s^{\bar{r}} \\
&+ \frac{1}{2} \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'}\left(\hat{\mathbf{G}}_{ia}^\Gamma\right)_q^{\bar{p}}\left(\hat{\mathbf{G}}_{jb}^{\Gamma'}\right)_s^{\bar{r}} \\
&- \Bigg(\sum_{(i,a)\in\Gamma} \left(\hat{\mathbf{G}}_{ia}^\Gamma\right)_q^{\bar{p}} \sum_{(j,b)\in\Gamma'} F_{I_j^{b,\Gamma'}}\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \Bigg)\left(\mathbf{G}_{I_0^{\Gamma'}}\right)_s^{\bar{r}} \\
&- \left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar{p}}\Bigg(\sum_{(j,b)\in\Gamma'} \left(\hat{\mathbf{G}}_{jb}^{\Gamma'}\right)_s^{\bar{r}} \sum_{(i,a)\in\Gamma} F_{I_i^{a,\Gamma'}}\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \Bigg) \Bigg\} \\
&+ \Bigg(-C_0\mathcal{F}_0^{\Gamma\Gamma'} + \sum_{\bar{\Gamma}\neq\Gamma,\Gamma'} \mathcal{F}_0^{\Gamma\Gamma'\bar{\Gamma}}\mathcal{L}^{\bar{\Gamma}} + \sum_{\substack{\bar{\bar{\Gamma}}>\bar{\Gamma} \\ \bar{\Gamma}\neq\Gamma,\Gamma' \\ \bar{\bar{\Gamma}}\neq\Gamma,\Gamma'}} \mathcal{F}_0^{\Gamma\Gamma'\bar{\Gamma}\bar{\bar{\Gamma}}}\mathcal{K}^{\bar{\Gamma}\bar{\bar{\Gamma}}} \Bigg)F_{I_0^\Gamma}F_{I_0^{\Gamma'}}\left(\mathbf{G}_{I_0^\Gamma}\right)_q^{\bar{p}}\left(\mathbf{G}_{I_0^{\Gamma'}}\right)_s^{\bar{r}} \Bigg\}
\end{aligned}
\tag{1.86}
$$

The notation $\otimes$ used in the above equations deserves some comments. It stands for the outer product, and in the above equations it always represents the outer product of two 2D matrices. For convenience, regular 2D matrices are put in bold face in the examples below, whereas its elements are in regular font. It results in a 4D object, according to the following expression:

$$
\left(A \otimes B\right)_{qs}^{pr} = A_q^p B_s^r .
\tag{1.87}
$$

That is, the entries of $\mathbf{A} \otimes \mathbf{B}$ are the products of all possible pairs of elements of the regular matrices $\mathbf{A}$ and $\mathbf{B}$. This is more easily visualised as a "matrix of matrices": a matrix of the same shape of $\mathbf{A}$,

where each entry is a matrix of the shape pf $\mathbf{B}$. For example, if $\mathbf{A}$ is 3×2 and $\mathbf{B}$ is 2×2:

$$
\mathbf{A} \otimes \mathbf{B} =
\begin{pmatrix}
\begin{pmatrix} A_1^1 B_1^1 & A_1^1 B_2^1 \\ A_1^1 B_1^2 & A_1^1 B_2^2 \end{pmatrix} & \begin{pmatrix} A_2^1 B_1^1 & A_2^1 B_2^1 \\ A_2^1 B_1^2 & A_2^1 B_2^2 \end{pmatrix} \\
\begin{pmatrix} A_1^2 B_1^1 & A_1^2 B_2^1 \\ A_1^2 B_1^2 & A_1^2 B_2^2 \end{pmatrix} & \begin{pmatrix} A_2^2 B_1^1 & A_2^2 B_2^1 \\ A_2^2 B_1^2 & A_2^2 B_2^2 \end{pmatrix} \\
\begin{pmatrix} A_1^3 B_1^1 & A_1^3 B_2^1 \\ A_1^3 B_1^2 & A_1^3 B_2^2 \end{pmatrix} & \begin{pmatrix} A_2^3 B_1^1 & A_2^3 B_2^1 \\ A_2^3 B_1^2 & A_2^3 B_2^2 \end{pmatrix}
\end{pmatrix}
=
\begin{pmatrix}
A_1^1 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix} & A_2^1 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix} \\
A_1^2 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix} & A_2^2 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix} \\
A_1^3 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix} & A_2^3 \begin{pmatrix} B_1^1 & B_2^1 \\ B_1^2 & B_2^2 \end{pmatrix}
\end{pmatrix}
=
\begin{pmatrix}
A_1^1 \mathbf{B} & A_2^1 \mathbf{B} \\
A_1^2 \mathbf{B} & A_2^2 \mathbf{B} \\
A_1^3 \mathbf{B} & A_2^3 \mathbf{B}
\end{pmatrix}
\tag{1.88}
$$

Hence, it is quite clear that $\mathbf{X}_\Gamma^{\Gamma'}$ is a "$K_\Gamma \times n^\Gamma$ matrix of $K_{\Gamma'} \times n^{\Gamma'}$ matrices", leading to a 4D object of shape $(K_\Gamma, n^\Gamma, K_{\Gamma'}, n^{\Gamma'})$. Some of its terms are indeed outer products of $K_\Gamma \times n^\Gamma$ matrices with $K_{\Gamma'} \times n^{\Gamma'}$ matrices. There are also some projections, carried out by the matrices $\Pi_{Y^\Gamma \perp}$. Whenever $\Pi_{Y^\Gamma \perp}$ acts over a regular matrix, the operation is an usual matrix multiplication, as in the expression for $\mathbf{C}_\Gamma$. However, in the expressions for $\mathbf{X}_\Gamma^{\Gamma'}$, projections might act on the "external big matrix", with indices pq, or over the "internal small matrices", with indices rs. The first case is done with $\Pi_{Y^\Gamma \perp} \otimes \mathbb{1}$, the second with $\mathbb{1} \otimes \Pi_{Y^{\Gamma'} \perp}$, whereas $\Pi_{Y^\Gamma \perp} \otimes \Pi_{Y^{\Gamma'} \perp}$ will carry projections at both levels. Note that the notation is totally compatible with the interpretation of "matrix of matrices": for example, if A is a 3×2 matrix of $K \times n$ matrices, $\mathbf{\Pi}$ is 3×3, and $\mathbf{\Pi}'$ is $K \times K$, then $\mathbf{\Pi} \otimes \mathbf{\Pi}'$ acting on A:

$$
\begin{aligned}
(\mathbf{\Pi} \otimes \mathbf{\Pi}')A &=
\begin{pmatrix}
\Pi_1^1 \mathbf{\Pi}' & \Pi_2^1 \mathbf{\Pi}' & \Pi_3^1 \mathbf{\Pi}' \\
\Pi_1^2 \mathbf{\Pi}' & \Pi_2^2 \mathbf{\Pi}' & \Pi_3^2 \mathbf{\Pi}' \\
\Pi_1^3 \mathbf{\Pi}' & \Pi_2^3 \mathbf{\Pi}' & \Pi_3^3 \mathbf{\Pi}'
\end{pmatrix}
\begin{pmatrix}
\mathbf{A}_1^1 & \mathbf{A}_2^1 \\
\mathbf{A}_1^2 & \mathbf{A}_2^2 \\
\mathbf{A}_1^3 & \mathbf{A}_2^3
\end{pmatrix} \\[2mm]
&=
\begin{pmatrix}
\Pi_1^1 \mathbf{\Pi}' \mathbf{A}_1^1 + \Pi_2^1 \mathbf{\Pi}' \mathbf{A}_1^2 + \Pi_3^1 \mathbf{\Pi}' \mathbf{A}_1^3 & \Pi_1^1 \mathbf{\Pi}' \mathbf{A}_2^1 + \Pi_2^1 \mathbf{\Pi}' \mathbf{A}_2^2 + \Pi_3^1 \mathbf{\Pi}' \mathbf{A}_2^3 \\
\Pi_1^2 \mathbf{\Pi}' \mathbf{A}_1^1 + \Pi_2^2 \mathbf{\Pi}' \mathbf{A}_1^2 + \Pi_3^2 \mathbf{\Pi}' \mathbf{A}_1^3 & \Pi_1^2 \mathbf{\Pi}' \mathbf{A}_2^1 + \Pi_2^2 \mathbf{\Pi}' \mathbf{A}_2^2 + \Pi_3^2 \mathbf{\Pi}' \mathbf{A}_2^3 \\
\Pi_1^3 \mathbf{\Pi}' \mathbf{A}_1^1 + \Pi_2^3 \mathbf{\Pi}' \mathbf{A}_1^2 + \Pi_3^3 \mathbf{\Pi}' \mathbf{A}_1^3 & \Pi_1^3 \mathbf{\Pi}' \mathbf{A}_2^1 + \Pi_2^3 \mathbf{\Pi}' \mathbf{A}_2^2 + \Pi_3^3 \mathbf{\Pi}' \mathbf{A}_2^3
\end{pmatrix} \\[2mm]
&= (\mathbf{\Pi} \otimes \mathbb{1})
\begin{pmatrix}
\mathbf{\Pi}' \mathbf{A}_1^1 & \mathbf{\Pi}' \mathbf{A}_2^1 \\
\mathbf{\Pi}' \mathbf{A}_1^2 & \mathbf{\Pi}' \mathbf{A}_2^2 \\
\mathbf{\Pi}' \mathbf{A}_1^3 & \mathbf{\Pi}' \mathbf{A}_2^3
\end{pmatrix} \\[2mm]
&= (\mathbb{1} \otimes \mathbf{\Pi}')
\begin{pmatrix}
\Pi_1^1 \mathbf{A}_1^1 + \Pi_2^1 \mathbf{A}_1^2 + \Pi_3^1 \mathbf{A}_1^3 & \Pi_1^1 \mathbf{A}_2^1 + \Pi_2^1 \mathbf{A}_2^2 + \Pi_3^1 \mathbf{A}_2^3 \\
\Pi_1^2 \mathbf{A}_1^1 + \Pi_2^2 \mathbf{A}_1^2 + \Pi_3^2 \mathbf{A}_1^3 & \Pi_1^2 \mathbf{A}_2^1 + \Pi_2^2 \mathbf{A}_2^2 + \Pi_3^2 \mathbf{A}_2^3 \\
\Pi_1^3 \mathbf{A}_1^1 + \Pi_2^3 \mathbf{A}_1^2 + \Pi_3^3 \mathbf{A}_1^3 & \Pi_1^3 \mathbf{A}_2^1 + \Pi_2^3 \mathbf{A}_2^2 + \Pi_3^3 \mathbf{A}_2^3
\end{pmatrix} .
\end{aligned}
\tag{1.89}
$$

Chapter 2

Hartree-Fock

History:

  2020 - Start

2.1 Notation

$\mathcal{H}$	The electronic Hamiltonian

2.2   Introduction

2.3   Equations

2.4   Flowchart of the algorithm

Chapter 3

Closed shell coupled cluster with singles and doubles

History:

2020 - Start

3.1 Notation

$\mathcal{H}$	The electronic Hamiltonian
g_{pqrs}	Two electron integrals over spatial orbitals
	$= \int dr_1 dr_2 \frac{\phi_p^*(r_1)\phi_r^*(r_2)\phi_q(r_1)\phi_s(r_2)}{r_{12}}$
L_{pqrs}	$= 2g_{pqrs} - g_{psrq}$
	(as suggested in [])

3.2 Introduction

The general equations of coupled cluster theory are "deceptively simple", as R. Bartlet has stated []:

$$E = \langle \Phi_0 | e^{-T} \mathcal{H} e^{T} | \Phi_0 \rangle \tag{3.1}$$

$$0 = \langle \Phi_\mu | e^{-T} \mathcal{H} e^{T} | \Phi_0 \rangle \tag{3.2}$$

where Φ_0 is the reference Slater determinant, $\mathcal{H}$ is the electronic Hamiltonian, and

$$T = \sum_\mu t_\mu \tau_\mu \tag{3.3}$$

is the cluster operator, with amplitudes t_μ associated to the excitations τ_μ. The amplitudes are obtained after solving equations (3.2), where $|\Phi_\mu\rangle = \tau_\mu |\Phi_0\rangle$ is the Slater determinant obtained after acting the excitation τ_μ on top of $|\Phi_0\rangle$.

These equations can be made specific to a particular case, and put in terms of the molecular integrals. This is done in Section 13.7 of reference [], that we followed closely in the present implementation. Below we show the main equations. Some of them are directly obtained from [], whereas in other cases we have adapted them to a form that is more directly related to our implementation. In section 3.4 we have a flowchart that show, as precisely as we could make, the connection between the equations and the code.

3.3 Equations

3.3.1 Energy

$$\begin{aligned} E &= \langle \Phi_0 | e^{-T} \mathcal{H} e^{T} | \Phi_0 \rangle \\ &= \langle \Phi_0 | \mathcal{H} e^{T} | \Phi_0 \rangle \\ &= E_{HF} + \sum_{aibj} (t_{ij}^{ab} + t_i^a t_j^b) L_{iajb} \end{aligned} \tag{3.4}$$

3.3.2 Residuals

Singles

$$\Omega_i^a = ... \tag{3.5}$$

Doubles

$$\Omega_{ij}^{ab} = ... \tag{3.6}$$

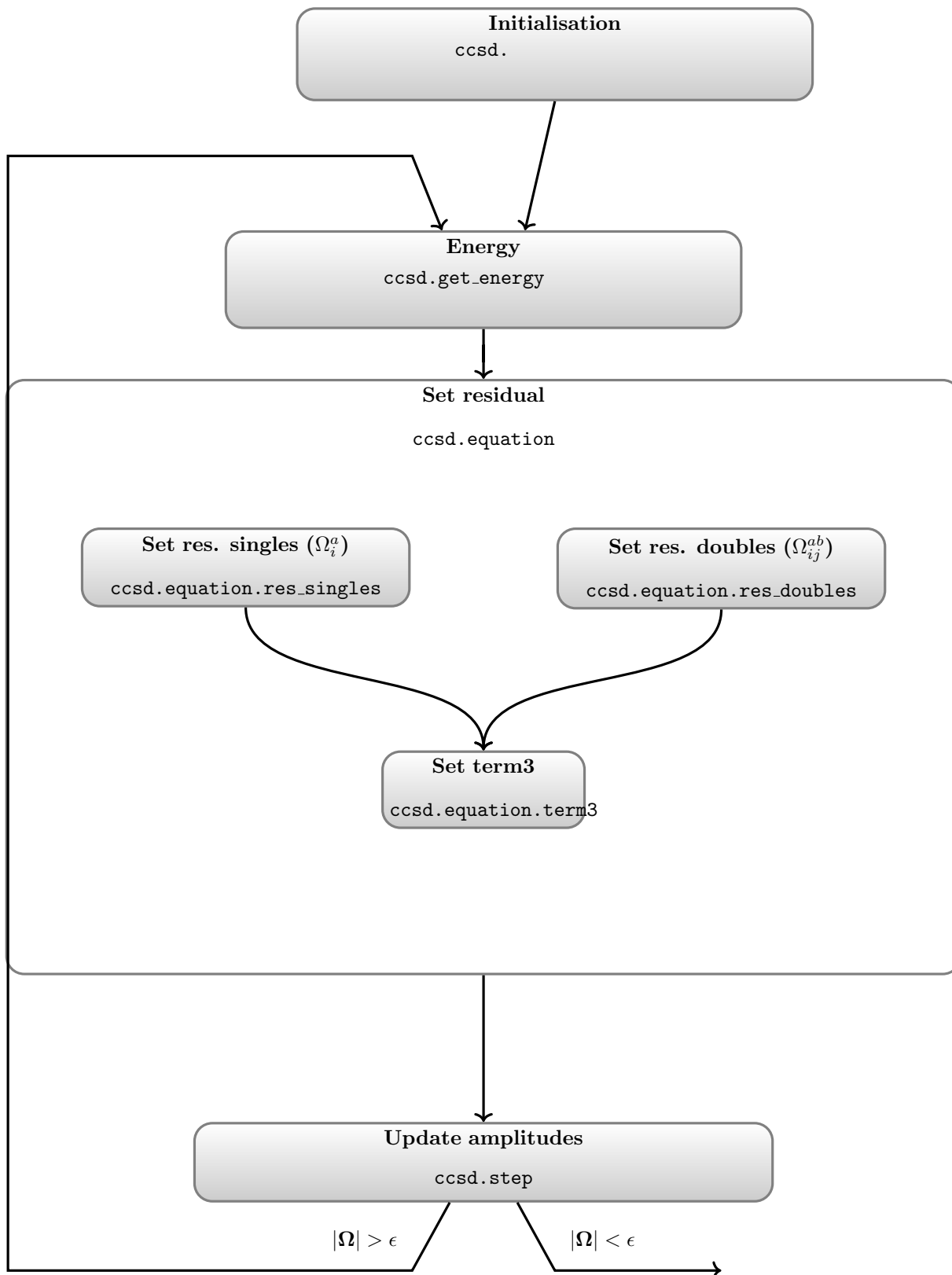### 3.3.3 Amplitudes update

$$t_\mu = t_\mu + \Delta t_\mu \tag{3.7}$$

$$\Delta t_i^a = \frac{\Omega_i^a(\mathbf{t})}{\epsilon_a - \epsilon_i} \tag{3.8}$$

$$\Delta t_{ij}^{ab} = \frac{\Omega_{ij}^{ab}(\mathbf{t})}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{3.9}$$

3.4   Flowchart of the algorithm

Chapter 4

The coupled-cluster manifold

History:

    2021 - Start

4.1 Notation

$\mathcal{H}$	The electronic Hamiltonian
Φ_0	The reference Slater determinant
$i, j, k, l, ...$	indices associated to occupied orbitals in Φ_0
$a, b, c, d, ...$	indices associated to virtual orbitals of Φ_0
Ψ	An arbitrary (correlated) wave function
Ψ_0	The FCI wave function associated of the ground state
Ψ_{CC}	An element of the CC-manifold:
	A coupled-cluster like wave function, not necessarily
	that one that solves the CC equations.
	(when the excitation level is implicit by the context
	or not important)
Ψ_{CCD}	An element of the CCD manifold
Ψ_{CCSD}	An element of the CCSD manifold

4.2   Introduction

In this chapter we describe the main equations used to study how the coupled-cluster manifold (the CC manifold) is related to an arbitrary wave function,[1] usually the exact ground state wave function. We receive as input this arbitrary wave function (Ψ_0) and we compare it with the manifold of coupled-cluster wave functions (Ψ_{CC}).

4.3   The distance in the intermediate normalisation

Let Φ_0 be our reference Slater determinant. Any state not orthogonal to Φ_0 can be uniquely written in the intermediate normalisation as:

$$\Psi = \Phi_0 + \sum_{I \neq I_0} c_I \Phi_I \,, \tag{4.1}$$

where Φ_I are excited Slater determinants (relative to Φ_0). The distance between two such wave functions is:

$$
\begin{aligned}
D_{\mathrm{IN}}(\Psi, \Psi') &= |\Psi - \Psi'| \\
&= \sqrt{\sum_{I \neq I_0} (c_I - c_I')^2}
\end{aligned}
\tag{4.2}
$$

4.4   To be in the right side of the CC manifold

Let Ψ_0 be an arbitrary wave function. We want to check if the CC manifold "curves towards it": We can see a CC manifold as obtained from the CI manifold (of corresponding type: CID $\leftrightarrow$ CCD, CISD $\leftrightarrow$ CCSD, etc) after "curving" it. It is expected that, being CC more accurate that CI, this "curving" goes towards the exact ground state wave function. This is what we want to investigate, and the precise definition is given below.

4.5   The CCD manifold

The CCD manifold is made by all wave functions of the type:

$$\Psi_{CCD} = e^{T_2} \Phi_0 = (1 + T_2 + \frac{1}{2!} T_2^2 + \frac{1}{3!} T_2^3 + ...) \Phi_0 \,, \tag{4.3}$$

where T_2 is defined by:

$$T_2 = \sum_{\substack{i<j \\ a<b}} t_{ij}^{ab} a_{ij}^{ab} \,. \tag{4.4}$$

Thus, the wave functions at the CCD manifold are linear combinations of the reference (Φ_0) determinant, of double excitations on top of Φ_0 (Φ_{ij}^{ab}), and higher-rank even excited determinants: quadruply

---

[1] These wave functions are not totally arbitrary: we will require them to be eigenfunctions of S_z. Except for this, they will be arbitrary FCI-like wave functions.

excited determinants, sextuply excited determinants, etc. For each one of these higher-rank excited determinant, its coefficient is given as a sum of products of t_{ij}^{ab}, that are the coefficients of the doubly excited determinants. Using Szabo's notation, these are

$$c_{ijkl}^{abcd} = t_{ij}^{ab} * t_{kl}^{cd} \,, \tag{4.5}$$

for quadruply excited determinants,

$$c_{ijklmn}^{abcdef} = t_{ij}^{ab} * t_{kl}^{cd} * t_{mn}^{ef} \,, \tag{4.6}$$

for sextuply excited determinants, etc.

The way that these are expressed in terms of t_{ij}^{ab} can be obtained by expanding the operator e^{T_2}. This has been done by Lehtola and coauthors []. They report in their paper the *cluster decomposition* of the FCI wave function, and show how each CI coefficient is decomposed in amplitudes. for the case with no truncation in the cluster operator T). Whenever we consider only the T_2 operator, we can simply neglect all contribution from amplitudes other than doubles. For instance:

$$
\begin{aligned}
c_{ijkl}^{abcd} = \quad & t_{ij}^{ab} t_{kl}^{cd} - t_{ij}^{ac} t_{kl}^{bd} + t_{ij}^{ad} t_{kl}^{bc} \\
& + t_{ij}^{bc} t_{kl}^{ad} - t_{ij}^{bd} t_{kl}^{ac} + t_{ij}^{cd} t_{kl}^{ab} \\
& - t_{ik}^{ab} t_{jl}^{cd} + t_{ik}^{ac} t_{jl}^{bd} - t_{ik}^{ad} t_{jl}^{bc} \\
& - t_{ik}^{bc} t_{jl}^{ad} + t_{ik}^{bd} t_{jl}^{ac} - t_{ik}^{cd} t_{jl}^{ab} \\
& + t_{il}^{ab} t_{jk}^{cd} - t_{il}^{ac} t_{jk}^{bd} + t_{il}^{ad} t_{jk}^{bc} \\
& + t_{il}^{bc} t_{jk}^{ad} - t_{il}^{bd} t_{jk}^{ac} + t_{il}^{cd} t_{jk}^{ab}
\end{aligned}
\tag{4.7}
$$

This rules for the decomposition until eight-rank excitations can be obtained from Lehtola's Github repository. The decomposition is done in `wave_functions.fci.cluster_decompose`. Note, however, that the above rules hold when we consider occupied orbitals in the reference coming before the virtuals (for each alpha or beta parts). If the orbitals are ordered in a different way a sign correction should be applied.

For each one of the higher-rank excited determinant (quadruples, etc), we say that the CCD manifold bends towards the FCI wave function in that direction if the sign of c_I in the FCI wave function is the same as of the corresponding cluster expansion in doubles, using the coefficients of doubles of the same wave function. For instance, if the FCI wave function is given in the intermediate normalization, the CCD manifold bends toward FCI in the direction of Φ_{ijkl}^{abcd} if:

$$\frac{c_{ij}^{ab} * c_{kl}^{cd}}{c_{ijkl}^{abcd}} > 0 \,. \tag{4.8}$$

Furthermore, the *vertical distance* between the wave function Ψ_0 and the CCD manifold is obtained as

$$D_{\text{IN}}(\Psi_0, \Psi_0^{CCD,vert}) \,, \tag{4.9}$$

where $\Psi_0^{CCD,vert}$ is the CCD wave function with the same double amplitudes as the coefficients of double

excitations as in Ψ_0:

$$D_{CCD}^{vert}(\Psi_0) = D_{IN}(\Psi_0, \Psi_0^{CCD,vert})$$

$$= \sqrt{\sum_i^a (c_i^a)^2 + \sum_{i<j<k}^{a<b<c} (c_{ijk}^{abc})^2 + \sum_{i<j<k<l}^{a<b<c<d} (c_{ijkl}^{abcd} - c_{ij}^{ab} * c_{kl}^{cd})^2 + \dots} \tag{4.10}$$

The calculation of the vertical distance between Ψ_0 and the CCD manifold, as well as the check if the manifold curves towards Ψ_0, is made at `fci.WaveFunctionFCI.compare_to_CC_manifold`.

4.6   The CCSD manifold

We will extend the ideas and equations discussed above to the CCSD manifold, formed by the wave functions of the form:

$$\Psi_{CCSD} = e^{T_1 + T_2} \Phi_0 = \left(1 + T_1 + T_2 + \frac{1}{2!}(T_1 + T_2)^2 + \frac{1}{3!}(T_1 + T_2)^3 + \dots \right) \Phi_0 \,, \tag{4.11}$$

where T_1 is further defined by:

$$T_1 = \sum_{i,a} t_i^a a_i^a \,. \tag{4.12}$$

Now the CC wave functions contain not only even rank excitations, but also singles, triples, etc. The decomposition triple and higher rank excitations in terms of single and doubles *amplitudes* can be done as before, following the work of Lehtola. However, now we cannot use the coefficients of double and single excitations in Ψ_0 directly, since the amplitudes of singles also contribute to the coefficient of doubly excited determinants in the CCSD wave function Thus, we have first to extract the contribution of singles from the coefficients of doubles in Ψ_0:

$$t_{ij}^{ab} = c_{ij}^{ab} - c_i^a * c_j^b \,. \tag{4.13}$$

Note that $c_i^a = t_i^a$.

Thus, we say that the CCSD manifold curves towards Ψ_0 in the direction of Φ_{ijk}^{abc} if

$$\frac{c_{ijk}^{abc}}{c_i^a * t_{jk}^{bc}} > 0 \,, \tag{4.14}$$

and that the CCSD manifold curves towards Ψ_0 in the direction of Φ_{ijkl}^{abcd} if

$$\frac{c_{ijkl}^{abcd}}{c_i^a * c_j^b * c_k^c * c_l^d + c_i^a * c_j^b * t_{kl}^{cd} + t_{ij}^{ab} * t_{kl}^{cd}} > 0 \,. \tag{4.15}$$

Finally the *vertical distance* between the wave function Ψ_0 and the CCSD manifold is obtained as $D_{IN}(\Psi_0, \Psi_0^{CCSD,vert})$, where $\Psi_0^{CCSD,vert}$ is the CCSD wave function with the same single amplitudes as the coefficients of single excitations as in Ψ_0, but doubles amplitudes given by Equation (4.13):

$$D_{CCSD}^{vert}(\Psi_0) = D_{IN}(\Psi_0, \Psi_0^{CCSD,vert})$$

$$= \sqrt{\sum_{i<j<k}^{a<b<c} (c_{ijk}^{abc} - c_i^a * t_{jk}^{bc})^2 + \sum_{i<j<k<l}^{a<b<c<d} (c_{ijkl}^{abcd} - c_i^a * c_j^b * c_k^c * c_l^d - c_i^a * c_j^b * t_{kl}^{cd} - t_{ij}^{ab} * t_{kl}^{cd})^2 + \dots}$$

$$\tag{4.16}$$