Notes on Fractional Occupation Number

B.Mussard
Laboratoire de Chimie Théorique and Institut du Calcul et de la Simulation
Université Pierre et Marie Curie, 75005 Paris, France
bastien.mussard@upmc.fr
http://www.lct.jussieu.fr/pagesperso/mussard/

14 décembre 2015

Table des matières

For the purpose of the analysis of the physics included in the diverse RPA correlation energies at our disposal, we investigate RPA calculations with fractional occupation numbers. These notes gather thoughts on the subject and can be very close to our implementation of these calculations in the `MOLPRO` package.

As much as possible, the indexes i, j, k, l will refer to fully occupied orbitals, p, q, r, s to partially occupied orbitals and a, b, c, d to unoccupied orbitals ; greek letters $\alpha, \beta, \gamma, \delta$ will refer to any *molecular* orbitals.

1 Fractional Occupation Number

1.1 SCF

For the purpose of fractional occupation number calculations, we need to modify the SCF and the RPA calculations. The SCF calculations have been generalized in `MOLPRO` through the modification of the computation of the density matrix, from the pre-existing integer occupation number situation, *i.e.* from :

$$P_{\mu\nu} = \sum_{i}^{\text{val.}} C_{i\mu} C_{i\nu}, \tag{1}$$

to :

$$P_{\mu\nu} = \sum_{i}^{\text{fully}} C_{i\mu} C_{i\nu} + \sum_{p}^{\text{partial}} n_p C_{p\mu} C_{p\nu}, \tag{2}$$

where we can choose to fractionally occupy any number of previously fully occupied orbitals, with $0 \leq n_p \leq 1$.

Note that the number of parameter used for the optimization is unchanged with respect to the $n_p = 1$ case, that is to say is $n_{\text{val.}}.n_{\text{unocc.}} = (n_{\text{fully}} + n_{\text{partial}}).n_{\text{unocc.}}$ This is how the generalization is described in the literature. It might or might not be the way it should be carried [EDIT : yes it is : the Gradient is null in other dimensions]. An example of this can be found in section 2.1.

1.2 RPA

We will consider three ways of calculating the RPA correlation energy : the DIAGO formulation, the PLASMON formulation, and the RCCD formulation. The DIAGO formulation is the brute diagonalisation of the RPA eigenvalue problem :

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & -\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} = \omega_n \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} \quad \text{and :} \quad E_c = \sum \omega_n - \sum A_{ia,ia}, \tag{3}$$

where the diagonalisation yields coupled positive and negative eigenvalues ω_n ; the summation is done on the positive eigenvalues.

The PLASMON formulation results from an effort to lower the dimension of the matrices to manipulate and reads :

$$\mathbf{M} = (\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B})^{1/2} \quad \text{and :} \quad E_c = \sum e_n^{1/2} - \sum A_{ia,ia}, \tag{4}$$

where the eigenvalues e_n of the half-dimensionned matrix $\mathbf{M}$ are the square of the eigenvalues of the RPA problem.

The RCCD formulation uses the Ricatti equations to compute amplitude and energy as follow :

$$\mathbf{B} + \mathbf{AT} + \mathbf{TA} + \mathbf{TBT} = \mathbf{0} \quad \text{and :} \quad E_c = \text{tr}\{\mathbf{BT}\} \tag{5}$$

The fractional occupation number version of the RPA calculations described by Yang *et al.* needs the following un-problematic change in the matrices $\mathbf{A}'$ and $\mathbf{B}$:

$$\text{from :} \quad A'_{ia,jb} = K_{ia,jb} - J_{ia,jb} \qquad \text{to :} \quad A'_{ia,jb} = \sqrt{n_i n_j (1 - n_a)(1 - n_b)}\,(K_{ia,jb} - J_{ia,jb}) \tag{6}$$

$$B_{ia,jb} = K_{ia,jb} - K_{ib,ja} \qquad B_{ia,jb} = \sqrt{n_i n_j (1 - n_a)(1 - n_b)}\,(K_{ia,jb} - K_{ib,ja}) \tag{7}$$

It also requires that the fractionally occupied orbitals be treated both as occupied and as virtual orbitals. This comes from the rigourous derivation using Green functions : the Green function is an ensemble average of N- and $N+1$-electrons systems Green functions ; the Fermi-level Heavyside function that appears in the Green functions, $\Theta(\epsilon_i - \epsilon_f)$, makes it so that the fractionally occupied orbitals are considered virtual in terms emerging from the N-electrons Green function and occupied in terms emerging from the $N+1$-electrons Green function.

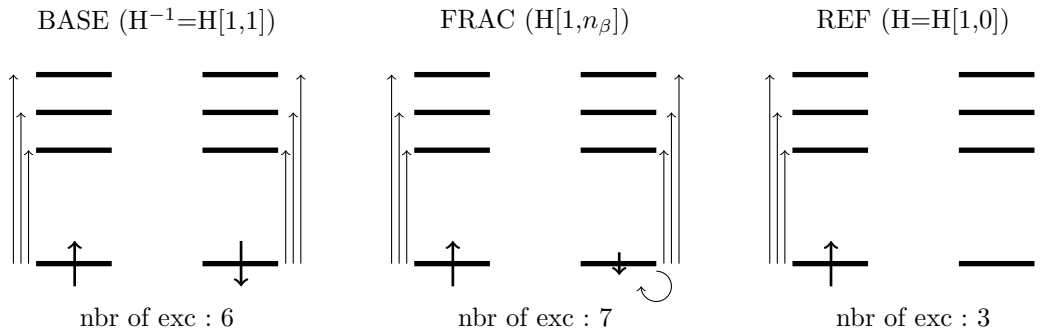
2 Number of excitations to consider

2.1 A closer look

Our fractional occupation number calculations are based on a "base" calculation and are compared to a "reference" calculation in the special cases where all fractional occupation numbers take an integer value.

This will be clear with the example of the study of the H[n_α,n_β] system, the hydrogen atom with $0 \leq n_\alpha \leq 1$ alpha-electron and $0 \leq n_\beta \leq 1$ beta-electron (the system H[1,0] is the hydrogen atom, H[1,1] is the H^- system). The "base" calculation *needs* to be the H^- system, with it's two fully occupied orbitals that may then be fractionally occupied.

What is shown in the sketch below is the theoretical number of *ia* excitations (*i.e.* the theoretical number of parameters of the orbital optimisation, *i.e.* the theoretical size of the RPA sub-matrices) for the "base" calculation, the fractional occupation calculation and the possible reference calculation, H[1,0]. The base calculation has 1*3+1*3=6 alpha+beta *ia* excitations, the H atom calculation (the reference) has 1*3=3 parameters. The fractional occupation calculation, with one fractionally occupied orbital, has 1*3+1*4=7 parameters.



The SCF fractional occupation number calculations, being a simple modification of the pre-existing routines, uses a number of 6 parameters for the optimisation. This implementation allows us to reproduce the results of [PRA 85 042507 (2012)] and [JCTC 5 786 (2009)]. One needs to understand that the H[1,0] system calculation coincides with the H atom, although the parameters of the optimization are redundant (6 instead of 3 for the H atom). Our RCCD implementation also reproduces the dRPA results of the PRA as well as the MP2 results of the JCTC.

2.2 Open question

The fractional occupation number calculations can be exotic cases, in the sense that the orbital energy differences can be null (at the position where a fractionally occupied orbital is taken as an occupied *and* as a virtual orbital : $\epsilon_{ii} = \epsilon_i - \epsilon_i = 0$) or, even, negative (in the case where two different fractionally occupied orbitals are taken as an occupied and a virtual orbital : $\epsilon_{ij} = \epsilon_i - \epsilon_j < 0$).

For this reason, the PLASMON formulation can be undoable ($(\mathbf{A}-\mathbf{B})^{1/2}$ tricky or impossible to compute). The RCCD and the DIAGO formulation pose no such problem.

2.3 Parametrization

Keep in mind from [London Mathematical Society Nonlinearity 27 1 (2014)] and [JCP 118 12 (2003)] that the optimization is usually done by rotating the orbitals with (in a picture where the orbitals are occupied or unoccupied) :

$$exp(\hat{\kappa}) \begin{pmatrix} \mathbf{I}_{oo} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} exp(-\hat{\kappa}) \qquad \text{with :} \quad \hat{\kappa} = \overbrace{\begin{pmatrix} \mathbf{0} & \text{-}\boldsymbol{\kappa}^*_{ou} \\ \boldsymbol{\kappa}_{ou} & \mathbf{0} \end{pmatrix}}^{\text{val.}\quad\text{unocc.}} \tag{8}$$

where the number of parameters is $n_{\text{val.}}.n_{\text{unocc.}}$. This is adapted in the literature, in a context where the orbitals are fully occupied, partially occupied or unoccupied, as :

$$exp(\hat{\kappa}) \begin{pmatrix} \mathbf{I}_{ff} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{pp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} exp(-\hat{\kappa}) \qquad \text{with :} \quad \hat{\kappa} = \overbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{\kappa}^*_{fu} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\kappa}^*_{pu} \\ \boldsymbol{\kappa}_{fu} & \boldsymbol{\kappa}_{pu} & \mathbf{0} \end{pmatrix}}^{\text{full}\quad\text{partial}\quad\text{unocc.}} \tag{9}$$

where the diagonal matrix $\mathbf{F}_{pp}$ contains the (fixed) fractional occupation numbers. Hence, the number of parameters for the SCF is unchanged, *i.e.* : $(n_{\text{fully}} + n_{\text{partial}}).n_{\text{unocc.}}$. We would like to know if this is correct and if we shouldn't actually perform calculations with the following scheme :

$$exp(\hat{\kappa}) \begin{pmatrix} \mathbf{I}_{ff} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{pp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} exp(-\hat{\kappa}) \qquad \text{with :} \quad \hat{\kappa} = \overbrace{\begin{pmatrix} \mathbf{0} & \boldsymbol{\kappa}^*_{fp} & \boldsymbol{\kappa}^*_{fu} \\ \boldsymbol{\kappa}_{fp} & \boldsymbol{\kappa}_{pp} & \boldsymbol{\kappa}^*_{pu} \\ \boldsymbol{\kappa}_{fu} & \boldsymbol{\kappa}_{pu} & \mathbf{0} \end{pmatrix}}^{\text{full}\quad\text{partial}\quad\text{unocc.}} \tag{10}$$

In this case the number of parameters would become $(n_{\text{fully}} + n_{\text{partial}}).(n_{\text{partial}} + n_{\text{unocc.}}) - n_{\text{partial}}$ ($\boldsymbol{\kappa}_{pp}$ does not have diagonal elements). We hence recover the dimensions of the fractional occupation number RPA problem (the matrices are actually of dimension $(n_{\text{fully}} + n_{\text{partial}}).(n_{\text{partial}} + n_{\text{unocc.}})$).

3 Gradient and Hessian

To investigate this, we want to derive the Gradient and Hessian of a fractional occupation number calculation The questions that we hope to elleviate concern both the dimension of the Hessian, and whether or not it undergoes the same transformation as described by Yang *et al.* (Eq. (7)).

Indeed, in practice, if we construct an Hessian from our RPA subroutines, it will have for the H system example a dimension of 7. Let's call this Hessian the "frac-dim-Hessian". This is not the Hessian that corresponds to the SCF calculation, which is of dimension 6 and will be called the "scf-dim-Hessian". This scf-dim-Hessian will be positive definite as it mesures the stability of the SCF calculation in the correct dimension space, but the frac-dim-Hessian could be non positive definite. In other words, in this case, the RPA calculation would be based (in terms of definition of matrices and applicability of the equations) on shaky basis, *i.e.* on an SCF calculation optimized in a lower dimension.

It is unclear if this issue is known of the litterature and if the PRA and JCTC take care of it. The fractional occupation number SCF calculations are always presented as a very simple modification from existing SCF (namely Eq (2)).

3.1 Usual case

We first derive the Gradient and Hessian in the classical, non-fractional occupation number, case.

3.1.1 Notations

Let us derive the usual energy :

$$E = \sum_i h_{ii} + \frac{1}{2}\sum_{ij}\langle ij||ij\rangle \tag{11}$$

$$f_{i\square} = h_{i\square} + \sum_j \langle ij||\square j\rangle \tag{12}$$

where the orbitals are parametrized with the following κ :

$$\hat{\kappa} = \sum_{ai} \kappa_{ai} a^\dagger i - \kappa_{ai}^* i^\dagger a \tag{13}$$

The non-vanishing terms of $\hat{\kappa}$ and $\hat{\kappa}^2$ (note that $\hat{\kappa}^\dagger = -\hat{\kappa}$) are :

$$\hat{\kappa}|i\rangle = \sum_a \kappa_{ai}|a\rangle \qquad \tfrac{1}{2}\hat{\kappa}^2|i\rangle = \tfrac{1}{2}\sum_{aj} -\kappa_{aj}^*\kappa_{ai}|j\rangle \tag{14}$$

$$\langle i|\hat{\kappa} = \sum_a \kappa_{ai}^*\langle a| \qquad \tfrac{1}{2}\langle i|\hat{\kappa}^2 = \tfrac{1}{2}\sum_{aj} -\kappa_{ai}^*\kappa_{aj}\langle j| \tag{15}$$

The updates of the bra and ket are as follow :

$$e^{\hat{\kappa}}|i\rangle = |i\rangle + \sum_a \kappa_{ai}|a\rangle - \tfrac{1}{2}\sum_{ja} \kappa_{aj}^*\kappa_{ai}|j\rangle$$
$$\langle i|e^{-\hat{\kappa}} = \langle i| + \sum_a \kappa_{ai}^*\langle a| - \tfrac{1}{2}\sum_{ja} \kappa_{ai}^*\kappa_{aj}\langle j|$$

3.1.2 First order

First order terms :

$$\Big(\sum_i h_{ii}\Big)^{(1)} = \sum_{i,a} \kappa_{ai}\langle i|h|a\rangle + \kappa_{ai}^*\langle a|h|i\rangle = \sum_{i,a}(\kappa_{ai} + \kappa_{ai}^*)h_{ai} \tag{16}$$

$$\Big(\tfrac{1}{2}\sum_{ij}\langle ij||ij\rangle\Big)^{(1)} = \tfrac{1}{2}\sum_{ij,a} \kappa_{ai}^*\langle aj||ij\rangle + \kappa_{aj}^*\langle ia||ij\rangle + \kappa_{ai}\langle ij||aj\rangle + \kappa_{aj}\langle ij||ia\rangle = \sum_{ij,a}(\kappa_{ai} + \kappa_{ai}^*)\langle aj||ij\rangle \tag{17}$$

Hence, first derivative of the energy (essentially selects the ia terms in the first order expressions above) :

$$\left.\frac{\partial E}{\partial \kappa_{ai}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = h_{ai} + \sum_j \langle aj||ij\rangle = f_{ai} \tag{18}$$

$$\left.\frac{\partial E}{\partial \kappa_{ai}^*}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = h_{ai} + \sum_j \langle aj||ij\rangle = f_{ai} \tag{19}$$

3.1.3 Second order

Second order terms :

$$\Big(\sum_i h_{ii}\Big)^{(2)} = \sum_{i,ab} \kappa_{bi}^*\kappa_{ai}\langle b|h|a\rangle - \sum_{ij,a} \kappa_{aj}^*\kappa_{ai}\langle j|h|i\rangle \tag{20}$$

$$\begin{aligned}\Big(\tfrac{1}{2}\sum_{ij}\langle ij||ij\rangle\Big)^{(2)} &= \tfrac{1}{2}\sum_{ij,ab} \kappa_{ai}^*\kappa_{bj}^*\langle ab||ij\rangle + \kappa_{ai}^*\kappa_{bi}\langle aj||bj\rangle + \kappa_{ai}^*\kappa_{bj}\langle aj||ib\rangle \\ &\quad + \kappa_{aj}^*\kappa_{bi}\langle ia||bj\rangle + \kappa_{aj}^*\kappa_{bj}\langle ia||ib\rangle + \kappa_{ai}\kappa_{bj}\langle ij||ab\rangle - \tfrac{1}{2}\sum_{ijk,a}(\kappa_{ak}^*\kappa_{aj} + \kappa_{aj}^*\kappa_{ak})\langle ij||ik\rangle \\ &= \sum_{ij,ab} \tfrac{1}{2}(\kappa_{ai}^*\kappa_{bj}^* + \kappa_{ai}\kappa_{bj})\langle ab||ij\rangle + (\kappa_{ai}^*\kappa_{bi})\langle aj||bj\rangle + (\kappa_{ai}^*\kappa_{bj})\langle aj||ib\rangle \\ &\quad - \tfrac{1}{2}\sum_{ijk,a}(\kappa_{ak}^*\kappa_{aj} + \kappa_{aj}^*\kappa_{ak})\langle ij||ik\rangle\end{aligned} \tag{21}$$

And second derivative of the energy (selects the ia and jb terms in the second order expressions) :

$$\left.\frac{\partial^2 E}{\partial \kappa_{ai}\kappa_{bj}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = \langle ij||ab\rangle \tag{22}$$

$$\left.\frac{\partial^2 E}{\partial \kappa^*_{ai}\kappa^*_{bj}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = \langle ij||ab\rangle \tag{23}$$

$$\begin{aligned}\left.\frac{\partial^2 E}{\partial \kappa_{ai}\kappa^*_{bj}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} &= \delta_{ij}h_{ab} - \delta_{ab}h_{ij} + \delta_{ij}\sum_k \langle bk||ak\rangle + \langle bi||ja\rangle + \delta_{ab}\sum_k \langle ik||jk\rangle \\ &= \delta_{ij}f_{ab} - \delta_{ab}f_{ij} + \langle bi||ja\rangle \end{aligned} \tag{24}$$

$$\begin{aligned}\left.\frac{\partial^2 E}{\partial \kappa^*_{ai}\kappa_{bj}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} &= \delta_{ij}h_{ab} - \delta_{ab}h_{ij} + \delta_{ij}\sum_k \langle ak||bk\rangle + \langle aj||ib\rangle + \delta_{ab}\sum_k \langle ik||jk\rangle \\ &= \text{idem} \end{aligned} \tag{25}$$

3.1.4 About κ

3.2 Frac case

3.2.1 Notations

We will now derive :

$$E = \sum_i h_{ii} + \sum_p n_p h_{pp} + \frac{1}{2}\sum_{ij}\langle ij||ij\rangle + \frac{1}{2}\sum_{ip} n_p\langle ip||ip\rangle + \frac{1}{2}\sum_{pq} n_p n_q \langle pq||pq\rangle \tag{26}$$

$$f_{i\square} = h_{i\square} + \sum_j \langle ij||\square j\rangle + \sum_q n_q\langle iq||\square q\rangle \tag{27}$$

$$f_{p\square} = n_p(h_{p\square} + \sum_j \langle pj||\square j\rangle + \sum_q n_q\langle pq||\square q\rangle) \tag{28}$$

where the orbitals are parametrized as follow :

$$\hat{\kappa} = \sum_{ai}\kappa_{ai}a^\dagger i - \kappa^*_{ai}i^\dagger a + \sum_{pi}\kappa_{pi}p^\dagger i - \kappa^*_{pi}i^\dagger p + \sum_{ap}\kappa_{ap}a^\dagger p - \kappa^*_{ap}p^\dagger a + \sum_{pq}\kappa_{pq}p^\dagger q \tag{29}$$

In the usual cases, part of the parametrization is left to freely chosen, since, for example, ...
For the case of the partial-partial block, we have :

$$\sum_{pq}\kappa_{pq}p^\dagger q = \sum_{p<q}\kappa_{pq}p^\dagger q + \sum_{p>q}\kappa_{pq}p^\dagger q = \sum_{p<q}\kappa_{pq}p^\dagger q + \kappa_{qp}q^\dagger p = \sum_{p<q}\kappa_{pq}p^\dagger q + \kappa^T_{pq}q^\dagger p \tag{30}$$

Note that the non-vanishing terms of $\hat{\kappa}^2$ are :

$$\hat{\kappa}^2|i\rangle = \sum_{ja} -\kappa^*_{aj}\kappa_{ai}|j\rangle + \sum_{pj} -\kappa^*_{pj}\kappa_{pi}|j\rangle + \sum_{ap} \kappa_{ap}\kappa_{pi}|a\rangle - \kappa^*_{ap}\kappa_{ai}|p\rangle + \sum_{pq} \kappa_{pq}\kappa_{qi}|p\rangle \tag{31}$$

$$\langle i|\hat{\kappa}^2 = \sum_{ja} -\kappa^*_{ai}\kappa_{aj}\langle j| + \sum_{ap} \kappa^*_{ap}\kappa^*_{pi}\langle a| - \kappa^*_{ai}\kappa_{ap}\langle p| + \sum_{pq} -\kappa^*_{qi}\kappa_{qp}\langle p| \tag{32}$$

$$\begin{aligned}\hat{\kappa}^2|p\rangle = &\sum_{ia} -\kappa^*_{pi}\kappa_{ai}|a\rangle - \kappa^*_{ai}\kappa_{ap}|i\rangle + \sum_{iq} -\kappa^*_{pi}\kappa_{qi}|q\rangle - \kappa^*_{qi}\kappa_{qp}|i\rangle \\ &+ \sum_{qa} \kappa_{aq}\kappa_{qp}|a\rangle - \kappa^*_{aq}\kappa_{ap}|q\rangle + \sum_{qr} \kappa_{rq}\kappa_{qp}|r\rangle\end{aligned} \tag{33}$$

$$\begin{aligned}\hat{\kappa}^2|p\rangle = &\sum_{ia} -\kappa^*_{ai}\kappa_{pi}\langle a| - \kappa^*_{ap}\kappa_{ai}\langle i| + \sum_{iq} -\kappa^*_{qi}\kappa_{pi}\langle q| + \kappa_{qi}\kappa_{pq}\langle i| \\ &+ \sum_{qa} -\kappa^*_{aq}\kappa_{pq}\langle a| - \kappa^*_{ap}\kappa_{aq}\langle q| + \sum_{qr} \kappa_{qr}\kappa_{pq}\langle r|\end{aligned} \tag{34}$$

The updates of the bra and ket are as follow :

$$\begin{aligned}|i\rangle \to &-\tfrac{1}{2}(\sum_{ja} \kappa^*_{aj}\kappa_{ai} + \sum_{pj} \kappa^*_{pj}\kappa_{pi})|j\rangle \\ &+ (\sum_{p} \kappa_{pi} - \tfrac{1}{2}\sum_{ap} \kappa^*_{ap}\kappa_{ai} + \tfrac{1}{2}\sum_{pq} \kappa_{pq}\kappa_{qi})|p\rangle \\ &+ (\sum_{a} \kappa_{ai} + \tfrac{1}{2}\sum_{ap} \kappa_{ap}\kappa_{pi})|a\rangle \\ \langle i| \to &-\tfrac{1}{2}(\sum_{ja} \kappa^*_{ai}\kappa_{aj} + \sum_{pj} \kappa^*_{pi}\kappa_{pj})\langle j| \\ &+ (\sum_{p} \kappa^*_{pi} - \tfrac{1}{2}\sum_{ap} \kappa^*_{ai}\kappa_{ap} - \tfrac{1}{2}\sum_{pq} \kappa^*_{qi}\kappa_{qp})\langle p| \\ &+ (\sum_{a} \kappa^*_{ai} + \tfrac{1}{2}\sum_{ap} \kappa^*_{ap}\kappa^*_{pi})\langle a|\end{aligned}$$

$$\begin{aligned}|p\rangle \to &-(\sum_{i} \kappa^*_{pi} + \tfrac{1}{2}\sum_{ia} \kappa^*_{ai}\kappa_{ap} + \tfrac{1}{2}\sum_{iq} \kappa^*_{qi}\kappa_{qp})|i\rangle \\ &+ (\sum_{q} \kappa_{qp} - \tfrac{1}{2}\sum_{qa} \kappa^*_{aq}\kappa_{ap} - \tfrac{1}{2}\sum_{iq} \kappa^*_{pi}\kappa_{qi} + \tfrac{1}{2}\sum_{qr} \kappa_{qr}\kappa_{rp})|q\rangle \\ &+ (\sum_{a} \kappa_{ap} - \tfrac{1}{2}\sum_{ia} \kappa^*_{pi}\kappa_{ai} + \tfrac{1}{2}\sum_{qa} \kappa_{aq}\kappa_{qp})|a\rangle \\ \langle p| \to &-(\sum_{i} \kappa_{pi} + \tfrac{1}{2}\sum_{ia} \kappa^*_{ap}\kappa_{ai} - \tfrac{1}{2}\sum_{iq} \kappa_{qi}\kappa_{pq})\langle i| \\ &+ (\sum_{q} \kappa_{pq} - \tfrac{1}{2}\sum_{qa} \kappa^*_{ap}\kappa_{aq} - \tfrac{1}{2}\sum_{iq} \kappa^*_{qi}\kappa_{pi} + \tfrac{1}{2}\sum_{qr} \kappa_{rq}\kappa_{pr})\langle q| \\ &+ (\sum_{a} \kappa^*_{ap} - \tfrac{1}{2}\sum_{ia} \kappa^*_{ai}\kappa_{pi} - \tfrac{1}{2}\sum_{qa} \kappa^*_{aq}\kappa_{pq})\langle a|\end{aligned}$$

3.2.2 First order

First order h_{ii} :

$$\begin{aligned}\Big(\sum_i h_{ii}\Big)^{(1)} &= \sum_{i,p} \kappa_{pi}\langle i|h|p\rangle + \sum_{i,a} \kappa_{ai}\langle i|h|a\rangle + \sum_{i,p} \kappa^*_{pi}\langle p|h|i\rangle + \sum_{i,a} \kappa^*_{ai}\langle a|h|i\rangle \\ &= \sum_{i,p} (\kappa_{pi} + \kappa^*_{pi})\langle i|h|p\rangle + \sum_{i,a} (\kappa_{ai} + \kappa^*_{ai})\langle i|h|a\rangle\end{aligned} \tag{35}$$

First order h_{pp} :

$$\Big(\sum_p n_p h_{pp}\Big)^{(1)} = \sum_{p,i} -n_p(\kappa_{pi}+\kappa^*_{pi})h_{ip} + \sum_{pq} n_p(\kappa_{pq}+\kappa_{qp})h_{pq} + \sum_{p,a} n_p(\kappa_{ap}+\kappa^*_{ap})h_{pa} \tag{36}$$

First order $\langle ij||ij\rangle$:

$$\Big(\tfrac{1}{2}\sum_{ij}\langle ij||ij\rangle\Big)^{(1)} = \sum_{ij,p}(\kappa_{pi}+\kappa^*_{pi})\langle pj||ij\rangle + \sum_{ij,a}(\kappa_{ai}+\kappa^*_{ai})\langle aj||ij\rangle \tag{37}$$

First order $\langle ip||ip\rangle$:

$$\begin{aligned}\Big(\sum_{ip} n_p\langle ip||ip\rangle\Big)^{(1)} &= \sum_{i,pq} n_p(\kappa_{qi}+\kappa^*_{qi})\langle qp||ip\rangle + \sum_{i,p,a} n_p(\kappa_{ai}+\kappa^*_{ai})\langle ap||ip\rangle - \sum_{ij,p} n_p(\kappa_{pj}+\kappa^*_{pj})\langle ij||ip\rangle \\ &\quad + \sum_{i,pq} n_p(\kappa_{qp}+\kappa_{pq})\langle iq||ip\rangle + \sum_{i,p,a} n_p(\kappa_{ap}+\kappa^*_{ap})\langle ia||ip\rangle \end{aligned}\tag{38}$$

First order $\langle pq||pq\rangle$:

$$\begin{aligned}\Big(\tfrac{1}{2}\sum_{pq} n_p n_q\langle pq||pq\rangle\Big)^{(1)} &= \sum_{pq,i} -n_p n_q(\kappa_{pi}+\kappa^*_{pi})\langle iq||pq\rangle + \sum_{pqr} n_p n_q(\kappa_{pr}+\kappa_{rp})\langle rq||pq\rangle \\ &\quad + \sum_{pq,a} n_p n_q(\kappa_{ap}+\kappa^*_{ap})\langle aq||pq\rangle \end{aligned}\tag{39}$$

Hence, derivative of energy :

$$\left.\frac{\partial E}{\partial \kappa_{ai}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = h_{ai} + \sum_j\langle aj||ij\rangle + \sum_p n_p\langle ap||ip\rangle = f_{ia} \tag{40}$$

$$\left.\frac{\partial E}{\partial \kappa^*_{ai}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = h_{ai} + \sum_j\langle aj||ij\rangle + \sum_p n_p\langle ap||ip\rangle = f_{ia} \tag{41}$$

$$\begin{aligned}\left.\frac{\partial E}{\partial \kappa_{pi}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} &= h_{ip} - n_p h_{ip} + \sum_j\langle pj||ij\rangle + \sum_q n_q\langle pq||iq\rangle - \sum_j n_p\langle ji||jp\rangle - \sum_q n_p n_q\langle iq||pq\rangle \\ &= h_{ip} + \sum_j\langle pj||ij\rangle + \sum_q n_q\langle pq||iq\rangle - n_p(h_{ip} + \sum_j\langle pj||ij\rangle + \sum_q n_q\langle pq||iq\rangle) \\ &= f_{ip} + f_{pi}\end{aligned}\tag{42}$$

$$\left.\frac{\partial E}{\partial \kappa^*_{pi}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = \text{idem} \tag{43}$$

$$\left.\frac{\partial E}{\partial \kappa_{ap}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = n_p(h_{pa} + \sum_i\langle ia||ip\rangle + \sum_q n_q\langle aq||pq\rangle) = f_{pa} \tag{44}$$

$$\left.\frac{\partial E}{\partial \kappa^*_{ap}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = \text{idem} \tag{45}$$

$$\begin{aligned}\left.\frac{\partial E}{\partial \kappa^*_{pq}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} &= n_p(h_{pq} + \sum_i\langle ip||iq\rangle + \sum_r n_r\langle pr||qr\rangle) + n_q(h_{pq} + \sum_i\langle ip||iq\rangle + \sum_r n_r\langle pr||qr\rangle) \\ &= f_{pq} + f_{qp}\end{aligned}\tag{46}$$

3.2.3 Automatic derivation

I wrote a Mathematica code to automatically derive all needed contributions (see #OTHER.nb#). The possibilities for superindexes formed from pairs of indexes of orbitals labelled fully-occupied (i,j,k,l), partially occupied (p,q,r,s) or unoccupied (a,b,c,d) are :

	i	p	a
i	s.o	$ip = pi^*$	$ia = ai^*$
p	pi	pq and pq^*	$pa = ap^*$
a	ai	ap	s.o

Hence, four types of unique superindexes are found : pi, ai, ap and pq (plus pi^*, ai^*, ap^* and pq^*). The possibilities for pairs of superindexes (found in derivations with respect to $\kappa_{\alpha\beta}\kappa_{\gamma\delta}$, with possible complex conjugate for one or both kappas) are as follow :

	pi	ai	ap	pq	...
pi^*	(1)				
ai^*	(2)	(5)			...
ap^*	(3)	(6)	(8)		
pq^*	(4)	(7)	(9)	(10)	
pi	(11)				
ai	(12)	(15)			...
ap	(13)	(16)	(18)		
pq	(14)	(17)	(19)	(20)	

where the first quarter will yield A matrices expressions, the second quarter B matrices, and the dots stand for complex conjugate of these expressions.

The following table attempts to summarize the results obtained. The factors in front of the bi-electronc integrals (which were checked to always be the correct ones) are shown, as well as, in the case of the A matrices, the factors in front of the fockian contributions.

A	indices	fac	fac
01	$pi^*.rj$	$(1-p)(1-r)$	$(1-p)$
02	$ai^*.rj$	$(1-r)$	0
03	$ap^*.rj$	$p(1-r)$	0
04	$pq^*.rj$	$(q-p)r$	(*1*)
05	$ai^*.bj$	1	1
06	$ap^*.bj$	p	(*2*)
07	$pq^*.bj$	$(q-p)$	0
08	$ap^*.br$	$p.r$	p
09	$pq^*.br$	$(q-p)r$	0
10	$pq^*.rs$	$(q-p)(s-r)$	$(q-p)$

B	indices	fac	fac
11	$pi.rj$	$(1-p)(1-r)$	0
12	$ai.rj$	$(1-r)$	0
13	$ap.rj$	$p(1-r)$	0
14	$pq.rj$	$(q-p)r$	(*1*)
15	$ai.bj$	1	0
16	$ap.bj$	p	0
17	$pq.bj$	$(q-p)$	0
18	$ap.br$	$p.r$	0
19	$pq.br$	$(q-p)r$	0
20	$pq.rs$	$(q-p)(s-r)$	(*3*)

The results show that we have :

$$\begin{aligned} A_{\alpha\beta,\gamma\delta} &= \langle\alpha\delta||\beta\delta\rangle(n_\beta - n_\alpha)(n_\delta - n_\gamma) + \Delta\epsilon_{\alpha\beta,\gamma\delta}(n_\beta - n_\alpha) \\ B_{\alpha\beta,\gamma\delta} &= \langle\alpha\gamma||\beta\delta\rangle(n_\beta - n_\alpha)(n_\delta - n_\gamma) \end{aligned} \tag{47}$$

3.2.4 YanMorCoh-JCP-2013

The derivation lead Yang *et al.* to go from (114) to (121) (pose $D_{ia} = \sqrt{n_i(1-n_a)}$, multiply by $\mathbf{D}_{ia}$ and use the diagonality of $\mathbf{1}$ and $\mathbf{\Delta}\epsilon$) :

$$\begin{aligned} \omega\mathbf{1}_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}^2_{jb}} &= \mathbf{\Delta}\epsilon_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}^2_{jb}} + \mathbf{A}_{ia,jb}\mathbf{X}_{jb} + \mathbf{B}_{ia,jb}\mathbf{Y}_{jb} \\ \omega\mathbf{D}_{ia}\mathbf{1}_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}^2_{jb}} &= \mathbf{D}_{ia}\mathbf{\Delta}\epsilon_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}^2_{jb}} + \mathbf{D}_{ia}\mathbf{A}_{ia,jb}\mathbf{D}_{jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}} + \mathbf{D}_{ia}\mathbf{B}_{ia,jb}\mathbf{D}_{jb}\frac{\mathbf{Y}_{jb}}{\mathbf{D}_{jb}} \\ \omega\mathbf{1}_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}} &= \Delta\epsilon_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}} + \mathbf{D}_{ia}\mathbf{A}_{ia,jb}\mathbf{D}_{jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}} + \mathbf{D}_{ia}\mathbf{B}_{ia,jb}\mathbf{D}_{jb}\frac{\mathbf{Y}_{jb}}{\mathbf{D}_{jb}} \\ \omega\begin{pmatrix}\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}\end{pmatrix}\begin{pmatrix}\tilde{\mathbf{X}} \\ \tilde{\mathbf{Y}}\end{pmatrix} &= \begin{pmatrix}\mathbf{\Delta}\epsilon + \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}} & \mathbf{\Delta}\epsilon + \tilde{\mathbf{A}}\end{pmatrix}\begin{pmatrix}\tilde{\mathbf{X}} \\ \tilde{\mathbf{Y}}\end{pmatrix} \end{aligned} \tag{48}$$

From the same (114), we can derive (multiply by $\mathbf{D}_{ia}^2\mathbf{D}_{jb}^2$) :

$$\begin{aligned}
\omega\mathbf{1}_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}^2} &= \mathbf{\Delta}\epsilon_{ia,jb}\frac{\mathbf{X}_{jb}}{\mathbf{D}_{jb}^2} + \mathbf{A}_{ia,jb}\mathbf{X}_{jb} + \mathbf{B}_{ia,jb}\mathbf{Y}_{jb} \\
\omega\mathbf{D}_{ia}^2\mathbf{1}_{ia,jb}\mathbf{X}_{jb} &= \mathbf{D}_{ia}^2\mathbf{\Delta}\epsilon_{ia,jb}\mathbf{X}_{jb} + \mathbf{D}_{ia}^2\mathbf{A}_{ia,jb}\mathbf{D}_{jb}^2\mathbf{X}_{jb} + \mathbf{D}_{ia}^2\mathbf{B}_{ia,jb}\mathbf{D}_{jb}^2\mathbf{Y}_{jb} \\
\omega\begin{pmatrix}\mathbf{D} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D}\end{pmatrix}\begin{pmatrix}\mathbf{X} \\ \mathbf{Y}\end{pmatrix} &= \begin{pmatrix}\tilde{\mathbf{\Delta}}\epsilon + \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}} & \tilde{\mathbf{\Delta}}\epsilon + \tilde{\mathbf{A}}\end{pmatrix}\begin{pmatrix}\mathbf{X} \\ \mathbf{Y}\end{pmatrix}
\end{aligned} \tag{49}$$

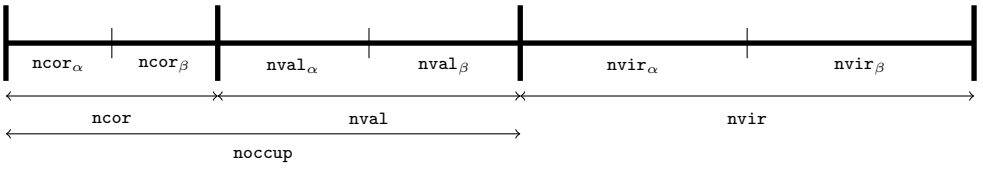
4 Structure of the open-shell RPA program

4.1 Up to now

In an open-shell context, a number `nAO` of atomic orbitals will yield a number `ntot`=2*`nAO` of molecular spin-orbitals. The orbital coefficients (and, from that, integrals, *etc*) are stored in a matrix :

$$\begin{array}{c} \\ \text{AO components} \end{array} \begin{array}{c} \text{MO orbitals} \\ \begin{pmatrix} \dots & \dots \\ & \end{pmatrix} \end{array}, \tag{50}$$

where the MO orbitals are sorted as follow :



Note that we always have `ncor`$_\alpha$=`ncor`$_\beta$, but it is possible that `nval`$_\alpha$≠`nval`$_\beta$ and, consequently, that `nvir`$_\alpha$≠`nvir`$_\beta$. Originally, and as a result of this structure, the super-indices occupied-virtual ia had the following succession of alpha/beta character : $\alpha\alpha$, $\alpha\beta$, $\beta\alpha$ and $\beta\beta$, and the integral matrices were as follow :

$$\mathbb{K} = \begin{array}{c} \\ \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \alpha\beta & \beta\alpha & \beta\beta \end{array} \\ \begin{pmatrix} \mathbf{K} & \mathbf{0} & \mathbf{0} & \mathbf{K} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K} & \mathbf{0} & \mathbf{0} & \mathbf{K} \end{pmatrix} \end{array} \quad \mathbb{K}' = \begin{array}{c} \\ \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \alpha\beta & \beta\alpha & \beta\beta \end{array} \\ \begin{pmatrix} \mathbf{K'} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K'} & \mathbf{0} \\ \mathbf{0} & \mathbf{K'} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K'} \end{pmatrix} \end{array} \quad \mathbb{J} = \begin{array}{c} \\ \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \alpha\beta & \beta\alpha & \beta\beta \end{array} \\ \begin{pmatrix} \mathbf{J} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J} \end{pmatrix} \end{array} \tag{51}$$

one can still use this version with the `NOSPINBLOCK` option. The default implementation now uses a reordering that allows to separate the matrices into non-spinflip and spinflip blocks, so that the calculations can be separated much like we separate the spin-adapted restricted calculations into singlet and triplet. Once reorganized, the matrices read as follow :

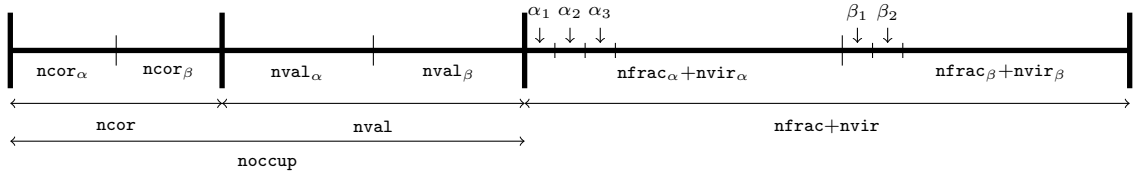
$$\mathbb{K} = \begin{array}{c} \\ \alpha\alpha \\ \beta\beta \\ \alpha\beta \\ \beta\alpha \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \beta\beta & \alpha\beta & \beta\alpha \end{array} \\ \begin{pmatrix} \mathbf{K} & \mathbf{K} & \mathbf{0} & \mathbf{0} \\ \mathbf{K} & \mathbf{K} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \end{array} \quad \mathbb{K}' = \begin{array}{c} \\ \alpha\alpha \\ \beta\beta \\ \alpha\beta \\ \beta\alpha \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \beta\beta & \alpha\beta & \beta\alpha \end{array} \\ \begin{pmatrix} \mathbf{K'} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K'} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K'} \\ \mathbf{0} & \mathbf{0} & \mathbf{K'} & \mathbf{0} \end{pmatrix} \end{array} \quad \mathbb{J} = \begin{array}{c} \\ \alpha\alpha \\ \beta\beta \\ \alpha\beta \\ \beta\alpha \end{array} \begin{array}{c} \begin{array}{cccc} \alpha\alpha & \beta\beta & \alpha\beta & \beta\alpha \end{array} \\ \begin{pmatrix} \mathbf{J} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J} \end{pmatrix} \end{array} \tag{52}$$

It is relatively easy to show that (as in the singlet/triplet case) only the RPAx-II flavor of RPA needs the spinflip block to be treated.

4.2 Fractional adaptation

Since, in our implementation, the fractionally occupied orbitals were originally fully occupied orbitals, only the virtual orbitals need some work to include the fractionally occupied ones.

This is now implemented in the following way : the general situation is that, from the 1 $\rightarrow$`noccup`$_\alpha$ alpha occupied orbitals, the user chooses a total of `nfrac`$_\alpha$ orbitals that are the $\alpha_1{}^{\text{th}}, \alpha_2{}^{\text{th}}, \ldots$ (idem for beta orbitals). The structure of the sorted orbitals in the program is then made as follow :



The variables are updated (`nvir`$\leftarrow$`nvir`+`nfrac`) so that, apart from the computation of the integrals, the code is mainly unchanged and seamlessly uses super-indices ia that run through fractionally occupied and virtual orbitals. The fractionally occupied orbitals $\alpha_1, \alpha_2, \alpha_3$ and β_1, β_2 are "repeated" in the virtual portion. This is achieved by the use of a dictionnary for the virtuals that reads $(\alpha_1, \alpha_2, \alpha_3, nvir_1, ..., nvir_\alpha)$ and $(\beta_1, \beta_2, nvir_1, ..., nvir_\beta)$ for the alpha and beta orbitals (the dictionnary reduces to $(nvir_1, ..., nvir_\alpha)$ and $(nvir_1, ..., nvir_\beta)$ in the usual case of integer occupation numbers).

New implementation with reduces dimension##

5 [ARCHIVE]

5.1 Basis for derivation

(Reproduce the beginning of "1.4 - CPHF" from JT and go to SuperCI)
With "usual and obvious" notations, we have :

$$|\kappa\rangle = e^{\hat{\kappa}}|0\rangle = \left(1 + \hat{\kappa} + \tfrac{1}{2}\hat{\kappa}^2 + \dots\right)|0\rangle, \quad \hat{\kappa} = \sum_{pq} \kappa_{pq} p^\dagger q = \sum_{'pq'} \kappa_{pq} p^\dagger q + \kappa_{qp} q^\dagger p \tag{53}$$

where $'pq'$ denotes unique pairs. When p and q are set to run on occupied and virtual, respectively, we obtain :

$$\hat{\kappa} = \sum_{ia} \kappa_{ai} a^\dagger i + \kappa_{ia} i^\dagger a \tag{54}$$

Imposing the anti-hermiticity of $\hat{\kappa}$ results in :

$$\hat{\kappa} = \sum_{ia} \kappa_{ai} a^\dagger i - \kappa^*_{ai} i^\dagger a \tag{55}$$

In the derivations, one will encounter $e^{\hat{\kappa}}|0\rangle$ as well as $\langle 0|e^{\hat{\kappa}^*}$. The adjoint operator $\hat{\kappa}^*$ can be written as (with and without the antihermicity imposed, *i.e.* from Eq. [54] or from Eq. [55]) :

$$\hat{\kappa}^* = \sum_{ia} \kappa^*_{ai} i^\dagger a + \kappa^*_{ia} a^\dagger i = -\left(\sum_{ia} \kappa_{ai} a^\dagger i - \kappa^*_{ai} i^\dagger a\right) = -\hat{\kappa} \tag{56}$$

From Eq. [55], we show that :

$$\hat{\kappa}^2 = \left(\sum_{ia} \kappa_{ai} a^\dagger i - \kappa^*_{ai} i^\dagger a\right)^2 = \sum_{iajb} \kappa_{ai}\kappa_{bj} a^\dagger i b^\dagger j - \sum_{iajb} \kappa_{ai}\kappa^*_{bj} a^\dagger i j^\dagger b - \sum_{iajb} \kappa^*_{ai}\kappa_{bj} i^\dagger a b^\dagger j + \sum_{iajb} \kappa^*_{ai}\kappa^*_{bj} i^\dagger a j^\dagger b \tag{57}$$

so that the action on the ket is :

$$\hat{\kappa}^2|0\rangle = \sum_{iajb} \kappa_{ai}\kappa_{bj}|^{ab}_{ij}\rangle - \sum_{ia} \kappa^*_{ai}\kappa_{ai}|0\rangle \tag{58}$$

and the action on the bra is :

$$\langle 0|\hat{\kappa}^{*2} = -\sum_{ia} \langle 0|\kappa^*_{ai}\kappa_{ai} + \sum_{iajb} \langle^{ab}_{ij}|\kappa^*_{ai}\kappa^*_{bj} \tag{59}$$

5.2 First and second derivatives

We have, in the case of a real $\boldsymbol{\kappa}$:

$$\begin{aligned} |\kappa\rangle &= \left(1 + \hat{\kappa} + \tfrac{1}{2}\hat{\kappa}^2\right)|0\rangle = |0\rangle + \kappa_{ia}|^a_i\rangle + \tfrac{1}{2}\kappa_{ia}\kappa_{jb}|^{ab}_{ij}\rangle - \tfrac{1}{2}\kappa_{ia}\kappa_{ia}|0\rangle \\ \langle\kappa| &= \langle 0|\left(1 + \hat{\kappa} + \tfrac{1}{2}\hat{\kappa}^2\right) = \langle 0| + \kappa_{ia}\langle^a_i| + \tfrac{1}{2}\kappa_{ia}\kappa_{jb}\langle^{ab}_{ij}| - \tfrac{1}{2}\kappa_{ia}\kappa_{ia}\langle 0| \end{aligned} \tag{60}$$

It is then easy to see that :

$$\left.\frac{\partial|\kappa\rangle}{\partial\kappa_{kc}}\right|_{\boldsymbol{\kappa}=0} = |_k^c\rangle \qquad \text{and :} \qquad \left.\frac{\partial\langle\kappa|}{\partial\kappa_{kc}}\right|_{\boldsymbol{\kappa}=0} = \langle_k^c|$$
$$\left.\frac{\partial^2|\kappa\rangle}{\partial\kappa_{kc}\kappa_{ld}}\right|_{\boldsymbol{\kappa}=0} = \begin{cases} |_{kl}^{cd}\rangle & \text{if } k\neq l \text{ or } c\neq d \\ -|0\rangle & \text{if } k=l \text{ and } c=d \end{cases} \qquad \text{and :} \qquad \left.\frac{\partial^2\langle\kappa|}{\partial\kappa_{kc}\kappa_{ld}}\right|_{\boldsymbol{\kappa}=0} = \begin{cases} \langle_{kl}^{cd}| & \text{if } k\neq l \text{ or } c\neq d \\ -\langle 0| & \text{if } k=l \text{ and } c=d \end{cases}$$
$$= (1-\delta_{kl}\delta_{cd})|_{kl}^{cd}\rangle - \delta_{kl}\delta_{cd}|0\rangle \qquad = (1-\delta_{kl}\delta_{cd})\langle_{kl}^{cd}| - \delta_{kl}\delta_{cd}\langle 0| \tag{61}$$

The Hartree-Fock energy and it's derivatives read :

$$E_{\mathrm{HF}} = \frac{\langle\kappa|\hat{H}|\kappa\rangle}{\langle\kappa|\kappa\rangle} = \frac{\langle 0|e^{-\hat{\kappa}}\hat{H}e^{\hat{\kappa}}|0\rangle}{\langle 0|e^{-\hat{\kappa}}e^{\hat{\kappa}}|0\rangle} = \langle 0|e^{-\hat{\kappa}}\hat{H}e^{\hat{\kappa}}|0\rangle \tag{62}$$
$$\left.\frac{\partial E_{\mathrm{HF}}}{\partial\kappa_{kc}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} = 2\langle 0|\hat{H}|_k^c\rangle \tag{63}$$
$$\begin{aligned}\left.\frac{\partial^2 E_{\mathrm{HF}}}{\partial\kappa_{kc}\partial\kappa_{ld}}\right|_{\boldsymbol{\kappa}=\mathbf{0}} &= 2\left(\frac{\partial^2\langle\kappa|}{\partial\kappa_{kc}\partial\kappa_{ld}}\hat{H}|\kappa\rangle + \frac{\partial\langle\kappa|}{\partial\kappa_{kc}}\hat{H}\frac{\partial|\kappa\rangle}{\partial\kappa_{ld}}\right) \\ &= 2\left((1-\delta_{kl}\delta_{cd})\langle_{kl}^{cd}|\hat{H}|0\rangle - \delta_{kl}\delta_{cd}\langle 0|\hat{H}|0\rangle + \langle_k^c|\hat{H}|_l^d\rangle\right)\end{aligned} \tag{64}$$

where the first term forms the matrix $\mathbf{B}$ and the two last terms the matrix $\mathbf{A}$.

5.3 SuperCI (cf CHAMP)

Let us write the wavefunction updated with the n parameters $\boldsymbol{\kappa}$ as a perturbation expansion :

$$|\kappa\rangle^{(1)} = |0\rangle + \sum_{j\geq 1}^{n}\frac{\partial|\kappa\rangle}{\partial\kappa_j}\kappa_j = \sum_{j\geq 0}^{n}|j\rangle\kappa_j \tag{65}$$

where $|j\rangle = \frac{\partial|\kappa\rangle}{\partial\kappa_j}$ and $\kappa_0 = 1$. The indices j can *a priori* refer to anything, in this case it will be obvious that they correspond to excitations $i\to a$. We wish to minimize the energy in the $n+1$ dimensions shown in the above sum :

$$\min_{\boldsymbol{\kappa}}\left\{\frac{\langle\kappa|\hat{H}|\kappa\rangle}{\langle\kappa|\kappa\rangle}\right\} \tag{66}$$

or, using a lagrangian framework with a lagrangian multiplier E :

$$\min_{\boldsymbol{\kappa}}\Big\{\langle\kappa|\hat{H}|\kappa\rangle - E\langle\kappa|\kappa\rangle\Big\} \tag{67}$$

i.e. we wish to solve the following derivative :

$$\frac{\partial}{\partial\kappa_i}\left\{\sum_{ij}\kappa_i\kappa_j\langle i|\hat{H}|j\rangle - E\sum_{ij}\kappa_i\kappa_j\langle i|j\rangle\right\} = 0 \quad\Leftrightarrow\quad \sum_j\kappa_j\langle i|\hat{H}|j\rangle = E\sum_j\kappa_j\langle i|j\rangle \quad\Leftrightarrow\quad \mathbf{H}\boldsymbol{\kappa} = E\mathbf{S}\boldsymbol{\kappa} \tag{68}$$

where the $(n+1).(n+1)$ matrix $\mathbf{H}$ and the $(n+1)$ vector $\boldsymbol{\kappa}$ read :

$$\mathbf{H} = \left(\begin{array}{c|ccc} \langle 0|\hat{H}|0\rangle & \langle 0|\hat{H}|j\rangle & \dots & \dots \\ \hline \langle j|\hat{H}|0\rangle & \langle i|\hat{H}|j\rangle & \dots & \dots \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & \end{array}\right) \quad \text{and :} \quad \boldsymbol{\kappa} = \begin{pmatrix} 1 \\ \kappa_1 \\ \vdots \\ \kappa_n \end{pmatrix} \tag{69}$$

(we can see here clearly that the indices j are excitations $i \to a$). Substracting E_{HF} from Eq. [68] gives :

$$\mathbf{H}\boldsymbol{\kappa} = (E - E_{\mathrm{HF}})\mathbf{S}\boldsymbol{\kappa} \tag{70}$$

where :

$$\mathbf{H} = \left(\begin{array}{c|ccc} 0 & \langle 0|\hat{H}|j\rangle & \dots & \dots \\ \hline \langle j|\hat{H}|0\rangle & & & \\ \vdots & & \mathbf{A} & \\ \vdots & & & \end{array}\right) \quad \text{and :} \quad \boldsymbol{\kappa} = \begin{pmatrix} 1 \\ \kappa_1 \\ \vdots \\ \kappa_n \end{pmatrix} \tag{71}$$

The eigenvector of $\mathbf{H}$ associated with it's lowest eigenvalue can be interpreted as the coefficients $\boldsymbol{\kappa}$ to be used to update the wavefunction.

5.4 SeegerPopleJCT1977

If we look at the derivation in the literature [Seeger,Pople JCP 66 3045 (1977)], we see :

$$\begin{aligned} &\kappa^*_{ia} A_{ia,jb} \kappa_{jb} + \frac{1}{2}\kappa_{ia} B^*_{ia,jb}\kappa_{jb} + \frac{1}{2}\kappa^*_{ia} B_{ia,jb}\kappa^*_{jb} \\ &= \frac{1}{2}\kappa^*_{ia} A_{ia,jb}\kappa_{jb} + \frac{1}{2}\kappa_{jb} A^*_{jb,ia}\kappa^*_{ia} + \frac{1}{2}\kappa_{ia} B^*_{ia,jb}\kappa_{jb} + \frac{1}{2}\kappa^*_{ia} B_{ia,jb}\kappa^*_{jb} \\ &= \begin{pmatrix} \kappa^* & \kappa \end{pmatrix} \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} \kappa \\ \kappa^* \end{pmatrix} \\ &= \begin{pmatrix} \kappa \\ \kappa^* \end{pmatrix}^* \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} \kappa \\ \kappa^* \end{pmatrix}, \end{aligned} \tag{72}$$

which makes me think that we have :

$$\begin{array}{c} \\ (ai) \\ (ia) \end{array} \begin{array}{c} \begin{array}{cc} (jb) & (bj) \end{array} \\ \begin{pmatrix} \dots & \dots \\ \dots & \dots \end{pmatrix} \end{array} = \begin{array}{c} \\ (ai) \\ -(ai)^* \end{array} \begin{array}{c} \begin{array}{cc} -(bj)^* & (bj) \end{array} \\ \begin{pmatrix} \frac{\partial -}{\partial \kappa_{ai}\kappa^*_{bj}}\Big|_{\boldsymbol{\kappa}=\mathbf{0}} & \frac{\partial +}{\partial \kappa_{ai}\kappa_{bj}}\Big|_{\boldsymbol{\kappa}=\mathbf{0}} \\ \frac{\partial +}{\partial \kappa^*_{ai}\kappa^*_{bj}}\Big|_{\boldsymbol{\kappa}=\mathbf{0}} & \frac{\partial -}{\partial \kappa^*_{ai}\kappa_{bj}}\Big|_{\boldsymbol{\kappa}=\mathbf{0}} \end{pmatrix} \end{array}, \tag{73}$$

which explains the minus sign that I get in the "A" part of my derivation.