Wavefunction-based Correlation Methods:

Basis rotations & Configuration Interaction (CI)

Zhendong Li

Beijing Normal University

April 15, 2021

Contents

Hubbard model

The Hubbard model proposed in 1963 (named after John Hubbard) is a simple model of interacting particles in a lattice in order to describe the transition between conducting and insulating systems (Mott insulators). Since then, it has been applied to the study of high-temperature superconductivity, quantum magnetism, and charge density waves.

The Fermi-Hubbard Hamiltonian reads

$$H = -t \sum_{<i,j>} \sum_{\sigma} (a^{\dagger}_{i\sigma} a_{j\sigma} + \text{h.c.}) + U \sum_{i} n_{i\alpha} n_{i\beta}, \tag{1}$$

where $a^{(\dagger)}_{i\sigma}$ is the creation (annihilation) operator for i-th site, t represents the hopping parameter, and U represents the on-site Coulomb repulsion.

Hubbard dimer

The two-site Hubbard model is a simple example for understanding electron correlation

$$H = -t \sum_{\sigma} (a_{1\sigma}^{\dagger} a_{2\sigma} + a_{2\sigma}^{\dagger} a_{1\sigma}) + U(n_{1\alpha} n_{1\beta} + n_{2\alpha} n_{2\beta}). \tag{2}$$

We focus on $N_e = K = 2$ (half-filling), where the Hilbert space is

$$\mathcal{H}_{N_e=2} = \mathrm{span}\{|1\bar{1}\rangle,\ |2\bar{2}\rangle,\ |1\bar{2}\rangle,\ |2\bar{1}\rangle\}. \tag{3}$$

By using the second quantization tool, we can find the Hamiltonian matrix (an exercise!) is

$$H = \begin{bmatrix} U & 0 & -t & -t \\ 0 & U & -t & -t \\ -t & -t & 0 & 0 \\ -t & -t & 0 & 0 \end{bmatrix} \tag{4}$$

which can be diagonalized using softwares like MATHEMATICA.

Spin and spatial symmetry adapted basis

Recall that $|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}(a^\dagger_{a\alpha}a_{i\alpha} \pm a^\dagger_{a\beta}a_{i\beta})|i\bar{i}\rangle = \frac{1}{\sqrt{2}}(|a\bar{i}\rangle \pm |i\bar{a}\rangle)$ is singlet (triplet) for $+$ $(-)$, we can define

$$|\Phi_c^{S+}\rangle = \frac{1}{\sqrt{2}}(|1\bar{2}\rangle + |2\bar{1}\rangle) \tag{5}$$

$$|\Phi_c^{T-}\rangle = \frac{1}{\sqrt{2}}(|1\bar{2}\rangle - |2\bar{1}\rangle) \tag{6}$$

The subscript 'c' means covalent, while the superscripts '$\pm$' represent the symmetric/antisymmetric with respect to the spatial inversion P_{12}, which exchanges sites 1 and 2.

We can easily check (H, S^2, S_z, P_{12}) forms a set of mutually commutating operators. $\mathcal{G} = \{I, P_{12}\}$ is a simple example of point group (discussed in later lectures). The advantage of introducing spin and spatial symmetry adapted basis is that the representation of H will become block-diagonal.

Spin and spatial symmetry adapted basis

Similarly, for the ionic configurations, we can introduce

$$|\Phi_i^{S+}\rangle = \frac{1}{\sqrt{2}}(|1\bar{1}\rangle + |2\bar{2}\rangle) \tag{7}$$

$$|\Phi_i^{S-}\rangle = \frac{1}{\sqrt{2}}(|1\bar{1}\rangle - |2\bar{2}\rangle) \tag{8}$$

Therefore, the Hilbert space can be decomposed as

$$\mathcal{H} = \mathcal{H}^{S+} \oplus \mathcal{H}^{S-} \oplus \mathcal{H}^{T-}, \tag{9}$$

$$\mathcal{H}^{S+} = \mathrm{span}\{|\Phi_c^{S+}\rangle,\ |\Phi_i^{S+}\rangle\}, \tag{10}$$

$$\mathcal{H}^{S-} = \mathrm{span}\{|\Phi_i^{S-}\rangle\}, \tag{11}$$

$$\mathcal{H}^{T-} = \mathrm{span}\{|\Phi_c^{T-}\rangle\}. \tag{12}$$

Spin and spatial symmetry adapted basis

The relation between the symmetry-adapted configurations with the original one is

$$(|\Phi_c^{S+}\rangle,\ |\Phi_i^{S+}\rangle,\ |\Phi_i^{S-}\rangle,\ |\Phi_c^{T-}\rangle) = (|1\bar{1}\rangle,\ |2\bar{2}\rangle,\ |1\bar{2}\rangle,\ |2\bar{1}\rangle)W, \tag{13}$$

$$W = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{14}$$

Consequently, the Hamiltonian in this basis $\tilde{H} = W^\dagger H W$ becomes block-diagonal,

$$\tilde{H} = \tilde{H}^{S+} \oplus \tilde{H}^{S-} \oplus \tilde{H}^{T-}, \tag{15}$$

$$\tilde{H}^{S+} = \begin{bmatrix} 0 & -2t \\ -2t & U \end{bmatrix}, \quad \tilde{H}^{S-} = [U], \quad \tilde{H}^{T-} = [0]. \tag{16}$$

Eigenvalues of Hubbard dimer

The eigenvalue of $\tilde{H}^{S+}$ satisfies $E = xt$

$$x^2 - Ux - 4 = 0, \quad x_\pm = \frac{U \pm \sqrt{U^2+16}}{2}. \tag{17}$$

The lowest eigenvalue is $x_-(U) = \frac{U-\sqrt{U^2+16}}{2}$, and the unnormalized eigenvector is $(\frac{1}{4}(U+\sqrt{16+U^2}), 1)$. The populations of the covalent and ionic components are

$$P_c(U) = \frac{1}{2}(1 + \frac{U}{\sqrt{U^2+16}}), \quad P_i(U) = \frac{1}{2}(1 - \frac{U}{\sqrt{U^2+16}}). \tag{18}$$

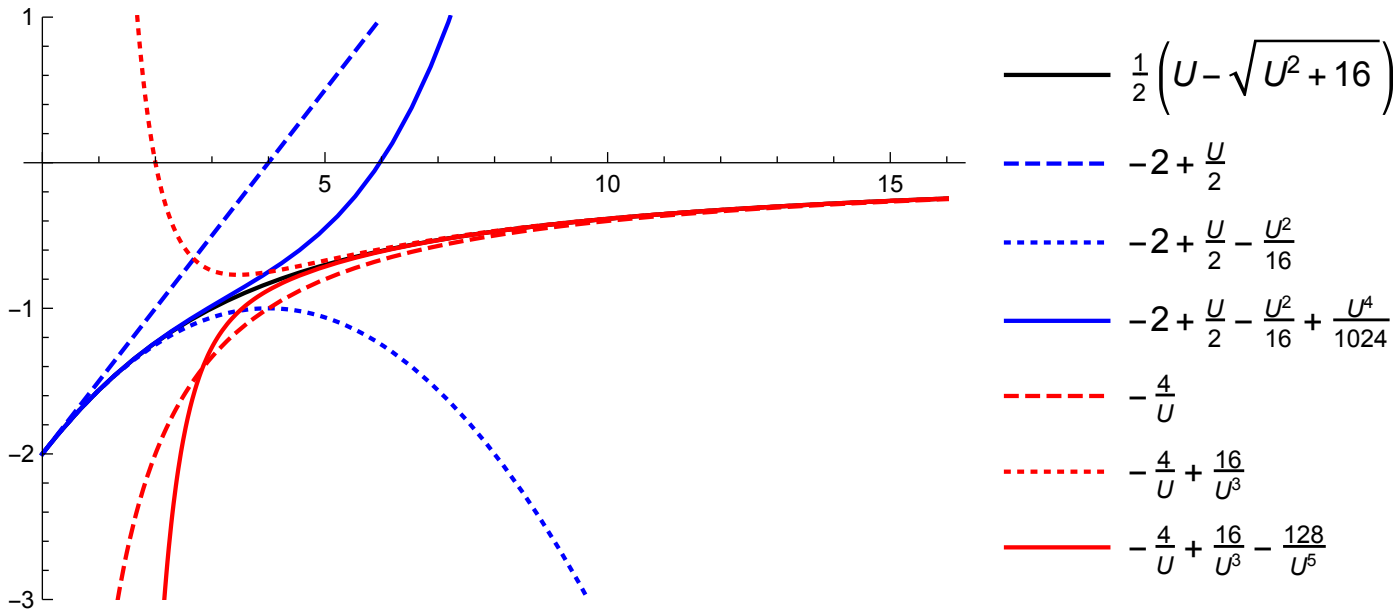
Therefore, the two limiting behaviors are

- Noninteracting limit $U = 0$: $P_c(U=0) = P_i(U=0) = \frac{1}{2}$;
- Strong interaction limit $U = \infty$: $P_c(U=\infty) = 1$ and $P_i(U=\infty) = 0$.

Eigenvalues of Hubbard dimer

The eigenvalue $x_-(U) = \frac{U-\sqrt{U^2+16}}{2}$ can be expanded with respect to U around these two limits.

- Small U: $x_-(U) \approx -2 + \frac{U}{2}$
- Large U: $x_-(U) \approx -\frac{4}{U}$



Molecular orbital basis

We can alteratively view the problem from a molecular orbital perspective.

$$(a_g^\dagger,\ a_u^\dagger) = (a_1^\dagger,\ a_2^\dagger)C, \quad C = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{19}$$

The configuration written in this basis is related with the previous one by

$$|\Phi_c^{S+}\rangle = \frac{1}{\sqrt{2}}(|1\bar{2}\rangle + |2\bar{1}\rangle) = \frac{1}{\sqrt{2}}(|g\bar{g}\rangle - |u\bar{u}\rangle) \tag{20}$$

$$|\Phi_i^{S+}\rangle = \frac{1}{\sqrt{2}}(|1\bar{1}\rangle + |2\bar{2}\rangle) = \frac{1}{\sqrt{2}}(|g\bar{g}\rangle + |u\bar{u}\rangle) \tag{21}$$

$$|\Phi_i^{S-}\rangle = \frac{1}{\sqrt{2}}(|1\bar{1}\rangle - |2\bar{2}\rangle) = \frac{1}{\sqrt{2}}(|g\bar{u}\rangle + |u\bar{g}\rangle) \tag{22}$$

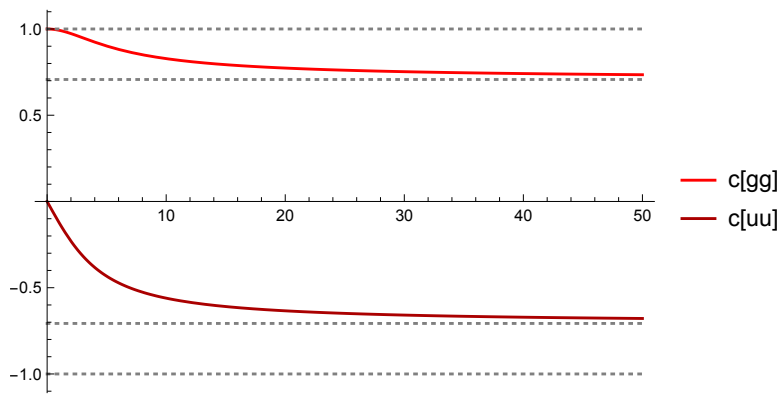
$$|\Phi_c^{T-}\rangle = \frac{1}{\sqrt{2}}(|1\bar{2}\rangle - |2\bar{1}\rangle) = \frac{1}{\sqrt{2}}(|g\bar{u}\rangle - |u\bar{g}\rangle) \tag{23}$$

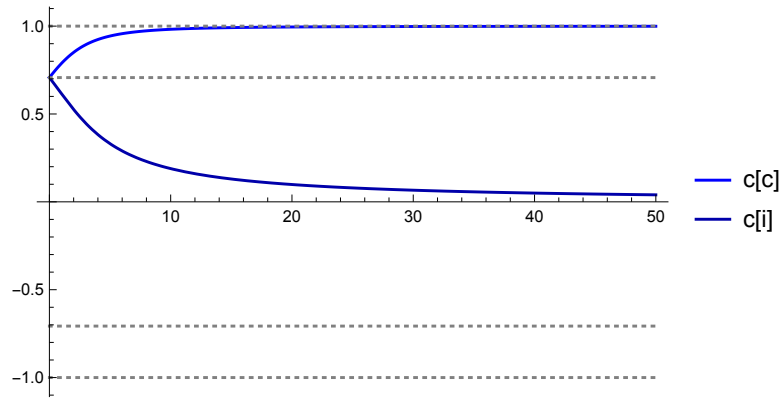
Molecular orbital basis

In the basis $(|g\bar{g}\rangle, |u\bar{u}\rangle)$, the Hamiltonian reads

$$\tilde{\tilde{H}}^{S+} = \begin{bmatrix} -2 + \frac{U}{2} & \frac{U}{2} \\ \frac{U}{2} & 2 + \frac{U}{2} \end{bmatrix} \tag{24}$$

Likewise, for the ground state, we can find the coefficients in this basis are





In summary, the two limits are described by

- Noninteracting limit: $\frac{1}{\sqrt{2}}(|\Phi_c^{S+}\rangle + |\Phi_i^{S+}\rangle) = |g\bar{g}\rangle$
- Strong correlation limit: $|\Phi_c^{S+}\rangle = \frac{1}{\sqrt{2}}(|g\bar{g}\rangle - |u\bar{u}\rangle)$

Contents

Basis rotations

We will show that Hartree-Fock (HF) theory, usually realized by solving the HF equation (stationary condition), can be viewed as an optimization problem, which can be solved by gradually rotating the orbitals to minimize $E = \langle \Phi(\theta) | \hat{H} | \Phi(\theta) \rangle$. This view is general and can be generalized to formulate CASSCF.

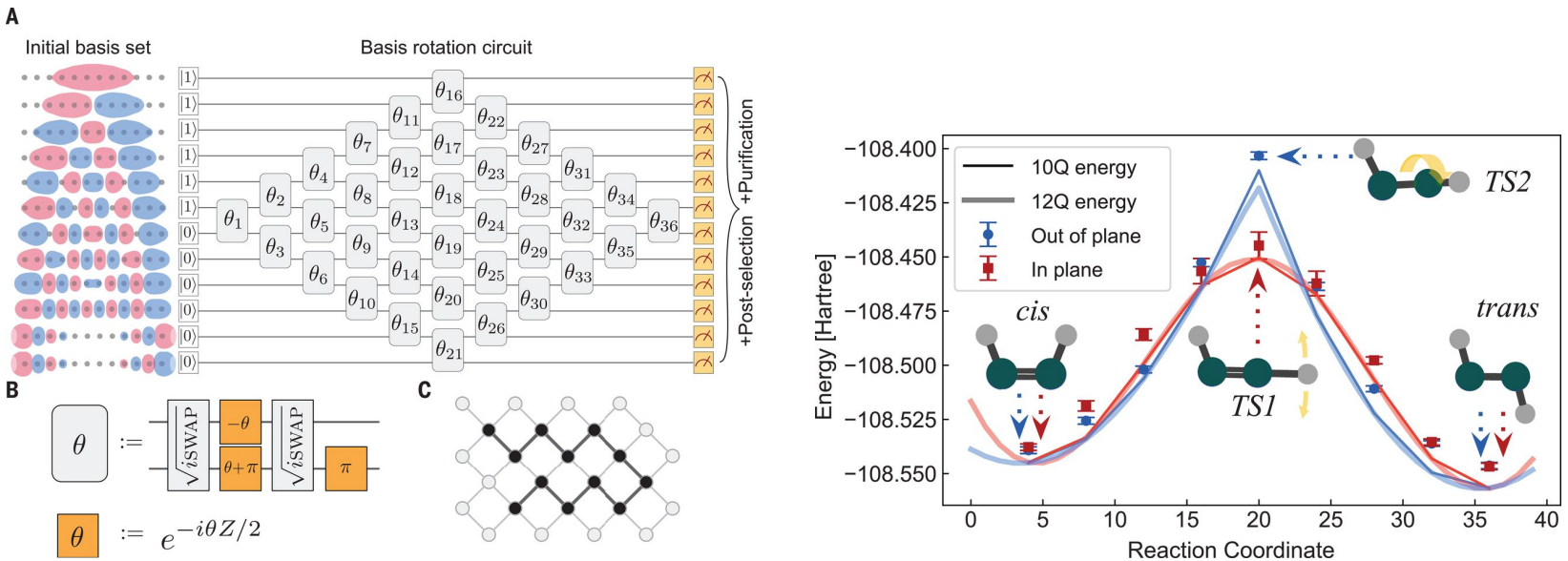


Figure 1: Hartree-Fock on a superconducting qubit quantum computer: Basis rotation circuit[1].

Basis rotations

Consider a set of spin orbitals $\{\tilde{\phi}_p\}$ obtained from another set $\{\phi_q\}$ by a unitary transformation:

$$|\tilde{\phi}_p\rangle = \sum_q |\phi_q\rangle U_{qp} \tag{25}$$

A unitary matrix $\mathbf{U}$ ($\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$) can be parameterized as

$$\mathbf{U} = e^{\boldsymbol{\kappa}}, \quad \boldsymbol{\kappa}^\dagger = -\boldsymbol{\kappa}, \tag{26}$$

where $\boldsymbol{\kappa}$ is an anti-Hermitian matrix.

Check (an exercise):

$$\mathbf{U}\mathbf{U}^\dagger = e^{\boldsymbol{\kappa}} e^{\boldsymbol{\kappa}^\dagger} = e^{\boldsymbol{\kappa}} e^{-\boldsymbol{\kappa}} = e^{\boldsymbol{\kappa}-\boldsymbol{\kappa}} = \mathbf{I} \tag{27}$$

Basis rotations

We demonstrate that the following unitary operator performs the orbital rotation.

Unitary operator for orbital rotations

$$\hat{U}(\boldsymbol{\kappa}) = e^{\hat{\kappa}}, \quad \hat{\kappa} = \sum_{pq} \kappa_{pq} a_p^\dagger a_q, \tag{28}$$

$$|\tilde{\phi}_p\rangle = \hat{U}(\boldsymbol{\kappa})|\phi_p\rangle = \sum_q |\phi_q\rangle [e^{\boldsymbol{\kappa}}]_{qp}. \tag{29}$$

Proof: It suffices to show that $\langle q|e^{-\hat{\kappa}}|p\rangle = [e^{-\boldsymbol{\kappa}}]_{qp}$. This can be verified term-by-term, e.g.,

$$\langle q|\hat{\kappa}^2|p\rangle = \langle q|\hat{\kappa}\hat{\kappa}|p\rangle = \sum_r \langle q|\hat{\kappa}|r\rangle \kappa_{rp} = \sum_r \kappa_{qr}\kappa_{rp} = [\boldsymbol{\kappa}^2]_{qp}. \tag{30}$$

$$\langle q|\hat{\kappa}^n|p\rangle = \langle q|\hat{\kappa}^{n-1}\hat{\kappa}|p\rangle = \sum_r \langle q|\hat{\kappa}^{n-1}|r\rangle \kappa_{rp} = \sum_r [\boldsymbol{\kappa}^{n-1}]_{qr}\kappa_{rp} = [\boldsymbol{\kappa}^n]_{qp}. \tag{31}$$

Basis rotations

The advantage of introducing the basis rotation operator $\hat{U}(\boldsymbol{\kappa})$ is that it can be used to transform not only the single particle states, but also many-particle basis. That is, $\hat{U}(\boldsymbol{\kappa})$ is an operator acting on the Fock space, describing the induced rotation by $\mathbf{U}$ at the single particle level!

Orbital rotations

$$\begin{aligned}
|\tilde{\Phi}\rangle &= \hat{U}(\boldsymbol{\kappa})|\Phi\rangle = \hat{U}(\boldsymbol{\kappa})a_1^\dagger \cdots a_N^\dagger|vac\rangle \\
&= \hat{U}(\boldsymbol{\kappa})a_1^\dagger[\hat{U}^\dagger(\boldsymbol{\kappa})\hat{U}(\boldsymbol{\kappa})]\cdots[\hat{U}^\dagger(\boldsymbol{\kappa})\hat{U}(\boldsymbol{\kappa})]a_N^\dagger[\hat{U}^\dagger(\boldsymbol{\kappa})\hat{U}(\boldsymbol{\kappa})]|vac\rangle \\
&= [\hat{U}(\boldsymbol{\kappa})a_1^\dagger\hat{U}^\dagger(\boldsymbol{\kappa})][\hat{U}(\boldsymbol{\kappa})\cdots\hat{U}^\dagger(\boldsymbol{\kappa})][\hat{U}(\boldsymbol{\kappa})a_N^\dagger\hat{U}^\dagger(\boldsymbol{\kappa})]\hat{U}(\boldsymbol{\kappa})|vac\rangle \\
&= \tilde{a}_1^\dagger \cdots \tilde{a}_N^\dagger|vac\rangle \qquad (32)
\end{aligned}$$

where $\tilde{a}_p^\dagger = \hat{U}(\boldsymbol{\kappa})a_p^\dagger\hat{U}^\dagger(\boldsymbol{\kappa})$. We will show that $\tilde{a}_p^\dagger = \sum_q a_q^\dagger U_{qp}$, which is consistent with the orbital rotation $|\tilde{\phi}_p\rangle = \sum_q |\phi_q\rangle U_{qp}$. The mathematical tool is called Baker-Campbell-Hausdorff (BCH) expansion.

Basis rotations

Baker-Campbell-Hausdorff (BCH) expansion

$$e^{A}Be^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \cdots \tag{33}$$

Proof: Let $f(x) = e^{xA}Be^{-xA}$, the derivative can be evaluated as

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = e^{xA}ABe^{-xA} - e^{xA}BAe^{-xA} = e^{xA}[A,B]e^{-xA}, \tag{34}$$

$$\frac{\mathrm{d}^2 f(x)}{\mathrm{d}^2 x} = e^{xA}[A,[A,B]]e^{-xA}, \tag{35}$$

Using $f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \frac{1}{3!}f'''(0)x^3 + \cdots$, then

$$e^{xA}Be^{-xA} = B + [A,B]x + \frac{1}{2}[A,[A,B]]x^2 + \frac{1}{3!}[A,[A,[A,B]]]x^3 + \cdots \tag{36}$$

Basis rotations

$$\tilde{a}_p^\dagger = e^{\hat{\kappa}} a_p^\dagger e^{-\hat{\kappa}} = a_p^\dagger + [\hat{\kappa}, a_p^\dagger] + \frac{1}{2!}[\hat{\kappa}, [\hat{\kappa}, a_p^\dagger]] + \frac{1}{3!}[\hat{\kappa}, [\hat{\kappa}, [\hat{\kappa}, a_p^\dagger]]] + \cdots \tag{37}$$

The commutators can be evaluated as

$$\begin{aligned} [\hat{\kappa}, a_p^\dagger] &= \sum_{rs} [\kappa_{rs} a_s^r, a^p] = \sum_{rs} \kappa_{rs} a^r \delta_s^p = \sum_r a^r \kappa_{rp}, \qquad (38) \\ [\hat{\kappa}, [\hat{\kappa}, a_p^\dagger]] &= \sum_r [\hat{\kappa}, a^r] \kappa_{rp} = \sum_r \left(\sum_s a^s \kappa_{sr} \right) \kappa_{rp} \\ &= \sum_s a^s \sum_r \kappa_{sr} \kappa_{rp} = \sum_s a^s [\boldsymbol{\kappa}^2]_{sp} \qquad (39) \end{aligned}$$

In summary,

$$\begin{aligned} \tilde{a}_p^\dagger &= a_p^\dagger + \sum_r a_r^\dagger \kappa_{rp} + \frac{1}{2!} \sum_r a_r^\dagger [\boldsymbol{\kappa}^2]_{rp} + \frac{1}{3!} \sum_r a_r^\dagger [\boldsymbol{\kappa}^3]_{rp} + \cdots \\ &= \sum_r a_r^\dagger [e^{\boldsymbol{\kappa}}]_{rp} = \sum_r a_r^\dagger U_{rp} \qquad (40) \end{aligned}$$

Contents

HF in terms of orbital rotations

We have shown that an arbitrary determinant can be parameterized as

$$|\Phi(\boldsymbol{\kappa})\rangle \triangleq \hat{U}(\boldsymbol{\kappa})|\Phi_0\rangle \tag{41}$$

In the Hartree-Fock approximation, the electronic wavefunction is approximated by a single Slater determinant, which is determined by variational optimization of the energy

$$E_{\mathrm{HF}} = \min_{\boldsymbol{\kappa}} E(\boldsymbol{\kappa}), \quad E(\boldsymbol{\kappa}) = \langle\Phi(\boldsymbol{\kappa})|\hat{H}|\Phi(\boldsymbol{\kappa})\rangle \tag{42}$$

where the HF energy function can be written as

$$\begin{aligned} E(\boldsymbol{\kappa}) &= \langle\Phi_0|e^{-\hat{\kappa}}\hat{H}e^{\hat{\kappa}}|\Phi_0\rangle \\ &= \langle\Phi_0|\hat{H}|\Phi_0\rangle + \langle\Phi_0|[\hat{H},\hat{\kappa}]|\Phi_0\rangle + \frac{1}{2!}\langle\Phi_0|[[\hat{H},\hat{\kappa}],\hat{\kappa}]|\Phi_0\rangle + \cdots \end{aligned} \tag{43}$$

HF in terms of orbital rotations

We can set the current expansion point to $\boldsymbol{\kappa} = 0$ by updating $|\Phi_0\rangle$. Then, the stationary condition implies that

$$\langle \Phi_0 | [\hat{H}, \delta\hat{\kappa}] | \Phi_0 \rangle = 0 \tag{44}$$

In general, an anti-Hermitian operator $\hat{\kappa}$ can be written more explicitly as

$$\begin{aligned}
\hat{\kappa} &= \sum_{pq} \kappa_{pq} a_p^\dagger a_q = \sum_{p>q} \kappa_{pq} a_p^\dagger a_q + \sum_{p<q} \kappa_{pq} a_p^\dagger a_q + \sum_p \kappa_{pp} a_p^\dagger a_p \\
&= \sum_{p>q} (\kappa_{pq} a_p^\dagger a_q - \kappa_{pq}^* a_q^\dagger a_p) + \sum_p \kappa_{pp} a_p^\dagger a_p \\
&= \sum_{p>q} \kappa_{pq}^R (a_p^\dagger a_q - a_q^\dagger a_p) + \sum_{p>q} \mathbb{i} \kappa_{pq}^I (a_p^\dagger a_q + a_q^\dagger a_p) + \sum_p \mathbb{i} \kappa_{pp}^I a_p^\dagger a_p
\end{aligned} \tag{45}$$

Given $|\Phi_0\rangle$, we can partition the orbitals into two sets: occupied (occ/closed: $i, j, k, \cdots$) and virtual (vir: $a, b, c, \cdots$).

HF in terms of orbital rotations

Examine the variations term-by-term:

$$\delta\hat{\kappa} = \sum_{p>q} \delta\kappa^R_{pq}(a^\dagger_p a_q - a^\dagger_q a_p) + \sum_{p>q} \mathbb{i}\delta\kappa^I_{pq}(a^\dagger_p a_q + a^\dagger_q a_p) + \sum_p \mathbb{i}\delta\kappa^I_{pp} a^\dagger_p a_p \quad (46)$$

- Term 3: $\langle\Phi_0|[\hat{H}, a^i_i]|\Phi_0\rangle = \langle\Phi_0|[\hat{H}, a^a_a]|\Phi_0\rangle = 0$, such that $\delta\kappa_{pp}$ is redundant, which corresponds to the variation of the phase of each orbital.
- Terms 1 and 2: since $a^j_i|\Phi_0\rangle = a^b_a|\Phi_0\rangle = 0$ for $i \neq j$ and $a \neq b$, then $\langle\Phi_0|[\hat{H}, a^j_i \pm a^i_j]|\Phi_0\rangle = \langle\Phi_0|[\hat{H}, a^b_a \pm a^b_a]|\Phi_0\rangle = 0$ This shows that the occ-occ and vir-vir rotations are redundant, since they do not change $|\Phi_0\rangle$.

$$\begin{aligned}
\langle\Phi_0|[\hat{H}, a^a_i - a^i_a]|\Phi_0\rangle &= \langle\Phi_0|\hat{H}a^a_i|\Phi_0\rangle + \langle\Phi_0|a^i_a\hat{H}|\Phi_0\rangle = f_{ia} + f_{ai} = 0 \\
\langle\Phi_0|[\hat{H}, a^a_i + a^i_a]|\Phi_0\rangle &= \langle\Phi_0|\hat{H}a^a_i|\Phi_0\rangle - \langle\Phi_0|a^i_a\hat{H}|\Phi_0\rangle = f_{ia} - f_{ai} = 0
\end{aligned}$$

which give $f_{ia} = f_{ai} = 0$ (so-called Brillouin's theorem).

HF in terms of orbital rotations

Therefore, $\hat{\kappa}$ for HF can be written as

$$\hat{\kappa} = \sum_{ai} \kappa_{ai}^{R}(a_a^\dagger a_i - a_i^\dagger a_a) + \sum_{ai} \mathbb{i}\kappa_{ai}^{I}(a_a^\dagger a_i + a_i^\dagger a_a) \tag{47}$$

This means that the $\boldsymbol{\kappa}$ matrix has the following structure

$$\boldsymbol{\kappa} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & \boldsymbol{\kappa}_{\mathrm{CV}} \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & -\boldsymbol{\kappa}_{\mathrm{VC}}^\dagger \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} \tag{48}$$

The number of nonredundant parameters: $2n_C n_V$ ($n_C n_V$) for complex (real) HF.

Extensions: For ROHF and CASSCF (discussed later), in which orbitals are partitioned into three sets, $\boldsymbol{\kappa}$ has the following structure

$$\boldsymbol{\kappa} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & \boldsymbol{\kappa}_{\mathrm{CO}} & \boldsymbol{\kappa}_{\mathrm{CV}} \\ \boldsymbol{\kappa}_{\mathrm{OC}} & \mathbf{0}_{\mathrm{OO}} & \boldsymbol{\kappa}_{\mathrm{OV}} \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \boldsymbol{\kappa}_{\mathrm{VO}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & -\boldsymbol{\kappa}_{\mathrm{OC}}^\dagger & -\boldsymbol{\kappa}_{\mathrm{VC}}^\dagger \\ \boldsymbol{\kappa}_{\mathrm{OC}} & \mathbf{0}_{\mathrm{OO}} & -\boldsymbol{\kappa}_{\mathrm{VO}}^\dagger \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \boldsymbol{\kappa}_{\mathrm{VO}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} \tag{49}$$

HF in terms of orbital rotations

Final remark: How is this formulation connected with the usual SCF formulation?

Recall the derivation of HF equation from variational principle

$$\delta L[\{\phi_i\}] = \delta[\langle\Phi_0|\hat{H}|\Phi_0\rangle - \sum_{ij}\epsilon_{ji}(\langle\phi_i|\phi_j\rangle - \delta_{ij})] \tag{50}$$

which leads to the HF equation

$$\hat{f}|\phi_i\rangle = \sum_j |\phi_j\rangle\epsilon_{ji} \tag{51}$$

This equation implies that $f_{ai} = \langle\phi_a|\hat{f}|\phi_i\rangle = \langle\phi_a|\sum_j|\phi_j\rangle\epsilon_{ji} = 0$. The canonical HF equation is obtained by choosing a special set of occupied orbitals, such that $\epsilon_{ji} = \langle\phi_j|\hat{f}|\phi_i\rangle$ is diagonalized, viz., $|\psi_i\rangle = \sum_j|\phi_j\rangle U_{ji}$ and $\boldsymbol{\epsilon} = \mathbf{U}\boldsymbol{\varepsilon}\mathbf{U}^\dagger$,

$$\hat{f}|\psi_i\rangle = |\psi_i\rangle\varepsilon_i \tag{52}$$

Full Configuration Interaction (FCI)

The configuration interaction (CI) is the earliest correlation method. The full CI (FCI) wavefunction can be written as:

$$\begin{aligned} |\Psi_{\mathrm{FCI}}\rangle &= |\Phi_{\mathrm{HF}}\rangle + \sum_{i,a} C_i^a \left|\Phi_i^a\right\rangle + \sum_{i<j,a<b} C_{ij}^{ab} \left|\Phi_{ij}^{ab}\right\rangle + \cdots \\ &= (1 + \hat{C}_1 + \hat{C}_2 + \cdots)|\Phi_{\mathrm{HF}}\rangle \\ &= (1 + \hat{C})|\Phi_{\mathrm{HF}}\rangle \end{aligned} \tag{53}$$

which is the exact method within a given one-electron basis. The coefficients (C_i^a, C_{ij}^{ab}, etc.) are linear variational parameters. The variational principle leads to a matrix eigenvalue problem

$$\mathbf{H}_{\mathrm{FCI}}\mathbf{C} = \mathbf{C}E \tag{54}$$

FCI

The FCI Hamiltonian matrix is

$$\mathbf{H}_{\mathrm{FCI}} = \begin{bmatrix} \langle\Phi_0|\hat{H}|\Phi_0\rangle & \langle\Phi_0|\hat{H}\left|\Phi_m^d\right\rangle & \langle\Phi_0|\hat{H}\left|\Phi_{mn}^{de}\right\rangle & \langle\Phi_0|\hat{H}\left|\Phi_{mnl}^{def}\right\rangle & \cdots \\ \langle\Phi_i^a|\hat{H}|\Phi_0\rangle & \langle\Phi_i^a|\hat{H}\left|\Phi_m^d\right\rangle & \langle\Phi_i^a|\hat{H}\left|\Phi_{mn}^{de}\right\rangle & \langle\Phi_i^a|\hat{H}\left|\Phi_{mnl}^{def}\right\rangle & \cdots \\ \left\langle\Phi_{ij}^{ab}\right|\hat{H}|\Phi_0\rangle & \left\langle\Phi_{ij}^{ab}\right|\hat{H}\left|\Phi_m^d\right\rangle & \left\langle\Phi_{ij}^{ab}\right|\hat{H}\left|\Phi_{mn}^{de}\right\rangle & \left\langle\Phi_{ij}^{ab}\right|\hat{H}\left|\Phi_{mnl}^{def}\right\rangle & \cdots \\ \left\langle\Phi_{ijk}^{abc}\right|\hat{H}|\Phi_0\rangle & \left\langle\Phi_{ijk}^{abc}\right|\hat{H}\left|\Phi_m^d\right\rangle & \left\langle\Phi_{ijk}^{abc}\right|\hat{H}\left|\Phi_{mn}^{de}\right\rangle & \left\langle\Phi_{ijk}^{abc}\right|\hat{H}\left|\Phi_{mnl}^{def}\right\rangle & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

The dimension of the matrix is

$$1 + C_N^1 C_{K-N}^1 + C_N^2 C_{K-N}^2 + \cdots + C_N^{\min(N,K-N)} C_{K-N}^{\min(N,K-N)} = C_K^N \tag{55}$$

The matrix elements can be evaluated using Slater-Condon rules or second quantization methods discussed before.

FCI

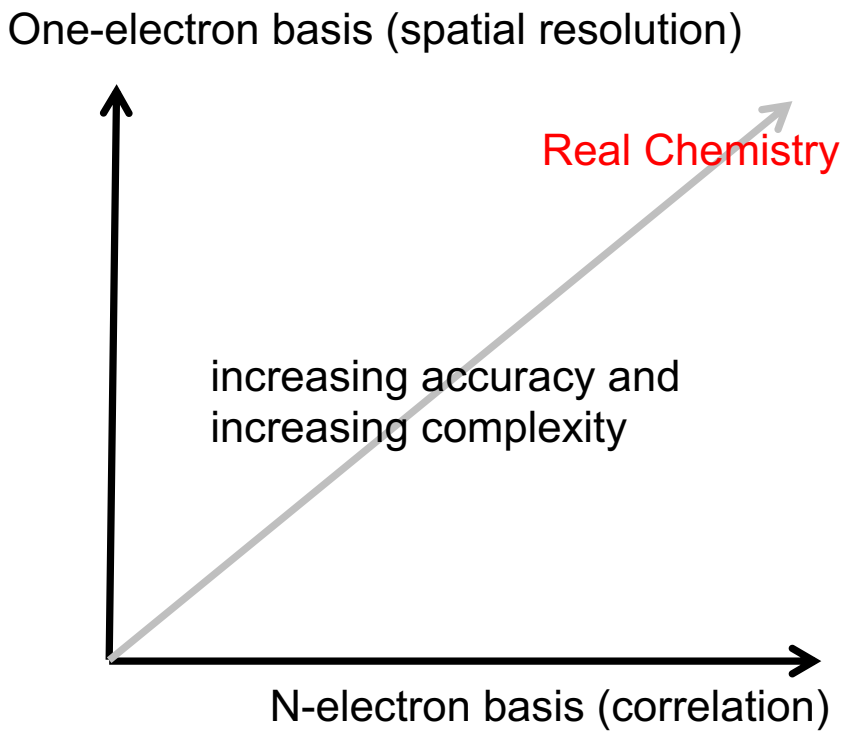
Using the property that $\hat{H}$ is two-body, as well as the Brillouin condition, $\mathbf{H}_{\mathrm{FCI}}$ based on the HF reference can be simplified to

$$\mathbf{H}_{\mathrm{FCI}} = \begin{bmatrix} \langle\Phi_0|\hat{H}|\Phi_0\rangle & 0 & \langle\Phi_0|\hat{H}|\Phi_{mn}^{de}\rangle & 0 & \cdots \\ 0 & \langle\Phi_i^a|\hat{H}|\Phi_m^d\rangle & \langle\Phi_i^a|\hat{H}|\Phi_{mn}^{de}\rangle & \langle\Phi_i^a|\hat{H}|\Phi_{mnl}^{def}\rangle & \cdots \\ \langle\Phi_{ij}^{ab}|\hat{H}|\Phi_0\rangle & \langle\Phi_{ij}^{ab}|\hat{H}|\Phi_m^d\rangle & \langle\Phi_{ij}^{ab}|\hat{H}|\Phi_{mn}^{de}\rangle & \langle\Phi_{ij}^{ab}|\hat{H}|\Phi_{mnl}^{def}\rangle & \cdots \\ 0 & \langle\Phi_{ijk}^{abc}|\hat{H}|\Phi_m^d\rangle & \langle\Phi_{ijk}^{abc}|\hat{H}|\Phi_{mn}^{de}\rangle & \langle\Phi_{ijk}^{abc}|\hat{H}|\Phi_{mnl}^{def}\rangle & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{56}$$

which is a highly sparse matrix with the number nonzero matrix elements per row goes as $O(N^2(K-N)^2)$. In comparison, its dimension C_K^N is exponential in the number of electrons N, as usually $K \sim O(N)$.

Truncated CI

FCI is only feasible for small systems! Rough estimation on the memory required to store a FCI vector: 2^{30} (8GB), 2^{40} (8TB), 2^{50} (8PB), etc. Therefore, we have to make some approximations in order to treat larger systems.



Truncated CI

- CIS (configuration interaction singles):

$$\mathbf{H}_{\mathrm{CIS}} = \begin{bmatrix} \langle \Phi_0 | \hat{H} | \Phi_0 \rangle & 0 \\ 0 & \langle \Phi_i^a | \hat{H} | \Phi_m^d \rangle \end{bmatrix} \tag{57}$$

  There is no improvement for the ground state if the orbitals are HF orbitals. Diagonalizing the block $[\langle \Phi_i^a | \hat{H} | \Phi_m^d \rangle]$ gives the simplest method for excited states.

- CISD (configuration interaction singles and doubles):

$$\mathbf{H}_{\mathrm{CISD}} = \begin{bmatrix} \langle \Phi_0 | \hat{H} | \Phi_0 \rangle & 0 & \langle \Phi_0 | \hat{H} | \Phi_{mn}^{de} \rangle \\ 0 & \langle \Phi_i^a | \hat{H} | \Phi_m^d \rangle & \langle \Phi_i^a | \hat{H} | \Phi_{mn}^{de} \rangle \\ \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_0 \rangle & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_m^d \rangle & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{mn}^{de} \rangle \end{bmatrix} \tag{58}$$

  This simple method is nowadays replaced by CCSD for several reasons.

Size-extensivity problem

Consider N noninteracting Hubbard dimer/hydrogen molecule in the minimal basis, the Hamiltonian can be written as

$$\hat{H} = \sum_{I=1}^{N} [\Delta b_I^\dagger b_I + V(b_I^\dagger + b_I)] \tag{59}$$

where $[b_I, b_J] = 0$ and $[b_I, b_J^\dagger] = \delta_{IJ}$. The exact ground state can be written as $|\Psi\rangle = \prod_{I=1}^{N} (|0_I\rangle + z|1_I\rangle)$. The exact can be written as

$$E_N^{\mathrm{FCI}} = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{N(1+z^*z)^{N-1}(\Delta z^*z + V(z^*+z))}{(1+z^*z)^N} \tag{60}$$

$$= \frac{N(\Delta z^*z + V(z^*+z))}{1+z^*z} = N E_{N=1}^{\mathrm{FCI}} \tag{61}$$

This property, called size-extensivity, is desirable for correlation methods.

Size-extensivity problem

The CISD energy function amounts to keeping terms up to z^*z in the denominator and numerator in Eq. (60), respectively,

$$E_N^{\mathrm{CISD}} = \frac{N(\Delta z^* z + V(z^* + z))}{(1 + N z^* z)} \tag{62}$$

The variational condition $\frac{\partial E_N^{\mathrm{CISD}}}{\partial z^*} = 0$ leads to $V + \Delta z - NVz^2 = 0$. In fact,

$$E_N^{\mathrm{FCI}} = N\frac{1}{2}(\Delta - \sqrt{\Delta^2 + 4V^2}), \quad E_N^{\mathrm{CISD}} = \frac{1}{2}(\Delta - \sqrt{\Delta^2 + 4NV^2}), \tag{63}$$

which agree with each other only for $N = 1$. For large N, $z \sim -1/\sqrt{N}$ and $E_N^{\mathrm{CISD}} \sim \sqrt{N}V$.

This example shows the lack of size-extensivity in CISD. We will introduce CCSD to solve this problem later, but we will not cover methods for approximately restoring size-extensivity (Davidson's correction, CEPA, etc., see Ref. [2]).

Contents

CASSCF

Complete-active-space SCF (CASSCF) wavefunction Ansatz[2]:

$$|\Psi_{\mathrm{CASSCF}}(\boldsymbol{\kappa}, \mathbf{S})\rangle = e^{\hat{\kappa}} e^{\hat{S}} |\Psi_0\rangle, \quad |\Psi_0\rangle = |\Psi_0^A\rangle |\Phi_0^C\rangle \tag{64}$$

where the orbital and configuration transformations are parameterized as

$$\hat{\kappa} = \sum_{pq} \kappa_{pq} a_p^\dagger a_q, \quad \boldsymbol{\kappa} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & \boldsymbol{\kappa}_{\mathrm{CO}} & \boldsymbol{\kappa}_{\mathrm{CV}} \\ \boldsymbol{\kappa}_{\mathrm{OC}} & \mathbf{0}_{\mathrm{OO}} & \boldsymbol{\kappa}_{\mathrm{OV}} \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \boldsymbol{\kappa}_{\mathrm{VO}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\mathrm{CC}} & -\boldsymbol{\kappa}_{\mathrm{OC}}^\dagger & -\boldsymbol{\kappa}_{\mathrm{VC}}^\dagger \\ \boldsymbol{\kappa}_{\mathrm{OC}} & \mathbf{0}_{\mathrm{OO}} & -\boldsymbol{\kappa}_{\mathrm{VO}}^\dagger \\ \boldsymbol{\kappa}_{\mathrm{VC}} & \boldsymbol{\kappa}_{\mathrm{VO}} & \mathbf{0}_{\mathrm{VV}} \end{bmatrix} \tag{65}$$

$$\hat{S} = \sum_{n>0} (S_{n0} |n\rangle\langle 0| - S_{0n}^* |0\rangle\langle n|), \quad \boldsymbol{S} = \begin{bmatrix} 0 & -S_1^* & -S_2^* & \cdots & -S_N^* \\ S_1 & 0 & 0 & \cdots & 0 \\ S_2 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_N & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{66}$$

CASSCF

The CASSCF energy function is

$$
\begin{aligned}
E(\boldsymbol{\kappa}, \mathbf{S}) &= \langle \Psi_{\mathrm{CASSCF}}(\boldsymbol{\kappa}, \mathbf{S}) | \hat{H} | \Psi_{\mathrm{CASSCF}}(\boldsymbol{\kappa}, \mathbf{S}) \rangle \\
&= \langle \Psi_0 | e^{-\hat{S}} e^{-\hat{\kappa}} \hat{H} e^{\hat{\kappa}} e^{\hat{S}} | \Psi_0 \rangle \\
&= \langle \Psi_0 | e^{-\hat{S}} (\hat{H} + [\hat{H}, \hat{\kappa}] + \frac{1}{2}[[\hat{H}, \hat{\kappa}], \hat{\kappa}] + \cdots) e^{\hat{S}} | \Psi_0 \rangle \\
&= \langle \Psi_0 | \hat{H} | \Psi_0 \rangle + \langle \Psi_0 | [\hat{H}, \hat{\kappa}] | \Psi_0 \rangle + \langle \Psi_0 | [\hat{H}, \hat{S}] | \Psi_0 \rangle \\
&+ \frac{1}{2}\langle \Psi_0 | [[\hat{H}, \hat{\kappa}], \hat{\kappa}] | \Psi_0 \rangle + \frac{1}{2}\langle \Psi_0 | [[\hat{H}, \hat{S}], \hat{S}] | \Psi_0 \rangle \\
&+ \langle \Psi_0 | [[\hat{H}, \hat{\kappa}], \hat{S}] | \Psi_0 \rangle + \cdots \\
&= E_0 + (\boldsymbol{g}_o^\dagger, \boldsymbol{g}_c^\dagger) \begin{pmatrix} \boldsymbol{x}_o \\ \boldsymbol{x}_c \end{pmatrix} + \frac{1}{2}(\boldsymbol{g}_o^\dagger, \boldsymbol{g}_c^\dagger) \begin{pmatrix} \mathbf{H}_{oo} & \mathbf{H}_{oc} \\ \mathbf{H}_{co} & \mathbf{H}_{cc} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_o \\ \boldsymbol{x}_c \end{pmatrix} + \cdots
\end{aligned} \tag{67}
$$

where $\boldsymbol{x}_o$ and $\boldsymbol{x}_c$ refers to the parameters for orbitals $\boldsymbol{\kappa}$ and configurations $\boldsymbol{S}$, respectively. This nonlinear function can be optimized by Newton's method introduced before.

MRCISD

Uncontracted MRCISD (uc-MRCISD): union of single and double excitations from each configuration - the problem is the number of configurations $\propto O(N_{\mathrm{CAS}})$.

Internally contraction (ic-MRCISD): suppose $|\Psi_0\rangle$ is the CASSCF reference

- Three classes of singly excited configurations:

$$|\Psi_i^x\rangle,\ |\Psi_x^a\rangle,\ |\Psi_i^a\rangle \tag{68}$$

- Eight classes of doubly excited configurations:

$$\begin{aligned} |\Psi_{ij}^{wx}\rangle,\ |\Psi_{iw}^{xy}\rangle \quad & \text{internal excitations,} \\ |\Psi_{ij}^{aw}\rangle,\ |\Psi_{iw}^{xa}\rangle,\ |\Psi_{wx}^{ya}\rangle \quad & \text{semi-internal excitations,} \\ |\Psi_{ij}^{ab}\rangle,\ |\Psi_{iw}^{ab}\rangle,\ |\Psi_{wx}^{ab}\rangle \quad & \text{double excitations,} \end{aligned} \tag{69}$$

  classified by the number of electrons excited in the external space.

Exercises

Problem 1: Find the eigenvalues of the following matrix for $s = +$ and $s = -$,

$$\begin{bmatrix} \Delta + K & K \\ sK & -(\Delta + K) \end{bmatrix}, \tag{70}$$

and discuss when the eigenvalues are not real.

Problem 2: Prove the following identity

$$[a_q^p, a_s^r] = \delta_q^r a_s^p - \delta_s^p a_q^r \tag{71}$$

Problem 3: If $\boldsymbol{\kappa} = -\mathbb{i}\theta\sigma_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ for real θ, show that

$$\mathbf{U} = e^{\boldsymbol{\kappa}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \cos(\theta) I_2 - \mathbb{i}\sin(\theta)\sigma_y \tag{72}$$

Problem 4: Let $|\Phi_0\rangle = |1\bar{1}\rangle$ and a double excitation operator $\hat{T} = \theta a_2^\dagger a_{\bar{2}}^\dagger a_{\bar{1}} a_1$ for real θ, show that

(i) $\hat{T}|\Phi_0\rangle = |2\bar{2}\rangle$;

(ii) $e^{\hat{T}}|\Phi_0\rangle = |1\bar{1}\rangle + \theta|2\bar{2}\rangle$;

(iii) $e^{\hat{T}-\hat{T}^\dagger}|\Phi_0\rangle = \cos(\theta)|1\bar{1}\rangle + \sin(\theta)|2\bar{2}\rangle$.

Reference I

-

[1] Google AI Quantum et al. "Hartree-Fock on a superconducting qubit quantum computer". In: *Science* 369.6507 (2020), pp. 1084–1089.

[2] Peter G Szalay et al. "Multiconfiguration self-consistent field and multireference configuration interaction methods and applications". In: *Chemical reviews* 112 (2012), pp. 108–181.